PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/00 — A2

(11) International Publication Number: **WO 99/07841**

(43) International Publication Date: 18 February 1999 (18.02.99)

(21) International Application Number: PCT/US98/16551

(22) International Filing Date: 7 August 1998 (07.08.98)

(30) Priority Data:
60/055,410 7 August 1997 (07.08.97) US

(71) Applicant *(for all designated States except US)*: WASHINGTON STATE UNIVERSITY RESEARCH FOUNDATION [US/US]; N.E. 1615 Eastgate Boulevard, Pullman, WA 99164-1802 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: OKITA, Thomas [US/US]; 910 S.W. Viento, Pullman, WA 99163 (US). GREENE, Thomas, W. [US/US]; 925 S.W. 82nd Terrace, Gainesville, FL 32607 (US). LAUGHLIN, Mary [US/US]; 2705 N.W. 115th Street, Vancouver, WA 98686 (US). KAHN, Michael [US/US]; 1440 N.W. Hall Drive, Pullman, WA 99163 (US). ITO, Hiroyuki [JP/US]; 1401 N.E. Merman Drive #705, Pullman, WA 99163 (US).

(74) Agent: STEPHENS, Donald, L., Jr.; Klarquist, Sparkman, Campbell, Leigh & Whinston, LLP, One World Trade Center, Suite 1600, 121 S.W. Salmon Street, Portland, OR 97204 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: REGULATORY MUTANTS OF ADP-GLUCOSE PYROPHOSPHORYLASE AND RELATED COMPOSITIONS AND METHODS

(57) Abstract

Up-regulated allosteric mutants of plant ADPG-PP enzymes having (1) higher sensitivity to allosteric activators (2) lower sensitivity to allosteric inhibitors; (3) increased starch production; (4) increased yield; (5) increased plant size; (6) increased growth rate and (7) increased number of seeds, and methods for efficiently producing and identifying such mutants.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

REGULATORY MUTANTS OF ADP-GLUCOSE PYROPHOSPHORYLASE AND RELATED COMPOSITIONS AND METHODS

FIELD OF THE INVENTION

This invention relates to starch biosynthesis in plants, particularly to regulatory mutants of ADPglucose pyrophosphorylase.

BACKGROUND

The regulation of starch biosynthesis. The first committed step in the synthesis of starch is the formation of ADPglucose, the substrate utilized by starch synthase. ADPglucose formation is catalyzed by the enzyme ADPglucose pyrophosphorylase (ADPG-PP). In many plant tissues, ADPG-PP is subject to allosteric control by small effector molecules, being activated by 3-phosphoglycerate (3-PGA) and inhibited by inorganic phosphate (Pi). It has been proposed that starch synthesis is controlled by the allosteric behavior of ADPG-PP (Preiss, In: *The Biochemistry of Plants*, Preiss, ed., New York: Academic Press, pp. 181-254, 1988; Preiss *et al.*, "Prospects for the production of cereals with improved starch properties," In: *Proceedings of the Improvement of Cereal Quality by Genetic Engineering*, Henry and Ronalds, eds., Plenum Press, 1994).

Several lines of evidence support the allosteric regulation of ADPG-PP as the dominant control of starch biosynthesis in leaf tissue. Several studies have evaluated the effect of Pi levels or the ratio of 3-PGA to Pi on the rate of starch biosynthesis in isolated chloroplasts or leaf discs (Heldt *et al.*, *Plant Physiol.* 59:1146-1155, 1977; Preiss, In: *Oxford Surveys of Plant Molecular and Cellular Biology*, Vol. 7, Miflin, ed., Oxford: Oxford Univ. Press, pp. 59-114, 1992; Preiss, In: *The Biochemistry of Plants*, Preiss, ed., New York:

Academic Press, pp. 181-254, 1988; Preiss, *Ann. Rev. Plant Physiol.* **33**:432-454, 1982; Preiss, *Biochem. Soc. Trans.* **19**:539-547, 1991; Preiss *et al.*, In: *Tailoring Genes for Crop Improvement: An Agricultural Perspective*, Kosuge *et al.*, eds., Plenum Press, pp. 133-152, 1987; Preiss *et al.*, In: *Biocatalysis in Agricultural Biotechnology*, Whitaker and Sonnet, eds., American Chemical Society, pp. 84-92, 1989). Theoretical modeling studies show a direct correlation between the relative levels of the activator 3-PGA in modulating ADPG-PP activity and, in turn, the rate of starch synthesis (Petterson and Ryde-Petterson, *J. Biochem* **179**:169-172, 1989). More recent studies on the biochemical responses of plant mutants that are defective in carbon metabolism also support a regulatory role for ADPG-PP in starch synthesis (Krukeberg *et al.*, *Biochem. J.* **261**:457-467, 1989; Neuhaus, *Planta* **178**:110-112, 1989). In particular, a *Chlamydomonas* starch mutant appears to contain an ADPG-PP that is defective in its activation by 3-PGA (Ball *et al.*, *Planta* **185**:17-26, 1991).

The structure of higher plant ADPG-PPs. The bacterial ADPG-PPs from both *Escherichia coli* and *Salmonella typhimurium* are encoded by a single gene locus, *glg*C (Preiss *et al.*, In: *Biocatalysis in Agricultural Biotechnology*, Whitaker and Sonnet, eds., American Chemical Society, pp. 84-92, 1989), which encodes a subunit of 48 kD that aggregates to form a homotetramer of 200 kD (Ghosh and Preiss, *J. Biol. Chem.* **241**:4491-4504, 1966). The higher plant enzyme is comprised of two distinct subunits, large subunit and small subunit, encoded by unique genes (Bae *et al.*, *Maydica* 35:317-322, 1990; Bhave *et al.*, *Plant Cell*

2:581-588, 1990; Copeland and Preiss, *Plant Physiol.* **68**:996-1001, 1981; Lin *et al.*, *Plant Physiol.* **99**:1175-1181, 1988; Lin *et al.*, *Plant Physiol.* **86**:1131-1135, 1988; Morell *et al.*, *Plant Physiol.* **85**:185-187, 1987; Okita *et al.*, *Plant Physiol.* **93**:785-790, 1990; Plaxton and Preiss, *Plant Physiol.* **83**:105-112, 1987; Smith-White and Preiss, *J. Mol. Evol.* **34**:449-464, 1992).

A number of genes that code for the large and small subunits have been isolated from a diverse group of plants (Ainsworth, *Plant Mol. Biol.* **23**:22-33, 1993, Anderson *et al.*, *J. Biol. Chem.* **264**:12238-12242, 1989; Anderson *et al.*, *Gene* **97**:199-205, 1991; Bae *et al.*, *Maydica* **35**:317-322, 1990; Bhave *et al.*, *Plant Cell* **2**:581-588, 1990; Jardin and Berhin, *Plant Mol. Biol.* **16**:349-351, 1991; La Cognata *et al.*, *Mol. Gen. Genet.* **246**:538-548, 1995; Muller-Rober *et al.*, *Mol. Gen. Genet.* **224**:136-146, 1990; Muller-Rober *et al.*, *Plant. Mol. Biol.* **27**:191-197, 1995; Nakata *et al.*, *J. Biol. Chem.* **269**:30798-30807, 1994; Nakata *et al.*, *Plant Mol. Biol.* **17**:1089-1093, 1991; Shaw and Hannah, *Plant Physiol.* **69**:1214, 1992; Smith-White and Preiss, *J. Mol. Evol.* **34**:449-464, 1992; Villand *et al.*, *Plant Mol. Biol.* **19**:381-389, 1992; Villand *et al.*, *Plant Physiol.* **100**:1617-1618, 1992; Villand *et al.*, *Plant Mol. Biol.* **23**:1279-1284, 1993). The large and small subunit sequences display about 30-55% amino acid identity. The primary sequences of the small subunits are more conserved between species (>90% identity) than the primary sequences between large and small subunits within a species (Nakata *et al.*, *Plant Mol. Biol.* 17:1089-1093, 1991; Smith-White and Preiss, *J. Mol. Evol.* **34**:449-464, 1992). This is consistent with immunological data that demonstrated that antibodies

raised against the spinach leaf small subunit but not to the large subunit cross-react with small subunits isolated from a number of different plants (Krishnan *et al.*, *Plant Physiol.* **81**:642-645, 1986; Okita *et al.*, *Plant Physiol.* **93**:785-790, 1990). These subunits are encoded by multiple genes that are expressed in specific tissues of the plant (Krishnan *et al.*, *Plant Physiol.* **81**:642-645, 1986; La Cognata *et al.*, *Mol. Gen. Genet.* **246**:538-548, 1995; Olive *et al.*, *Plant Mol. Biol.* **12**:525-538, 1989; Smith-White and Preiss, *J. Mol. Evol.* **34**:449-464, 1992; Villand *et al.*, *Plant Mol. Biol.* **19**:381-389, 1992; Villand *et al.*, *Plant Physiol.* **100**:1617-1618, 1992; Villand *et al.*, *Plant Mol. Biol.* **23**:1279-1284, 1993). In potato, three large subunit sequences have been isolated that differ in their spatial and temporal patterns of expression (La Cognata *et al.*, *Mol. Gen. Genet.* **246**:538-548, 1995). Two of the genes are expressed in leaves and tubers, while a third is expressed solely in tubers. The three large subunit sequences share from 68% to 75% identity at the amino acid level.

Structure-function relationships of the plant ADP-glucose pyrophosphorylase. Using chemical labeling and photoaffinity labeling strategies, three residues in the bacterial enzyme have been identified to be important in binding of the allosteric effectors or substrates. Lys_{39} is located near the activator binding site (Parsons and Preiss, *J. Biol. Chem.*, **253**:7638-7645, 1978), Lys_{195} participates in binding of the substrate glucose-1-phosphate (Glc 1-P) (Hill *et al.*, *J. Biol. Chem.* **266**:12455-12460, 1991), while Tyr_{114} is located at or near the ATP/ADPglucose binding site (Lee and Preiss, *J. Biol. Chem.* **261**:1058-1064, 1986). Lys_{39} and Lys_{195} are conserved in all small subunits of the plant enzyme, while a Phe is present in the corresponding Tyr_{114}

position in the plant's large subunit. The conservation of the two Lys residues suggest that they play a similar role in the plant enzymes. This has been confirmed by replacement of the corresponding Lys_{195} residue in the potato small subunit with a glutamate by site-directed mutagenesis experiments (Preiss and Sivak, "Starch synthesis in sinks and sources," In: *Photoassimilate Distribution in Plants and Crops: Source-Sink Relationships*, Zamski and Schaffer, eds., Marcel Dekker, Inc., 1995). This mutation increases the binding constant (K_m) of the enzyme for glucose-1-phosphate from 80 μM to over 45 mM without any effect on the kinetic parameters for the other substrates.

The role of specific amino-acid residues for allosteric regulation and catalysis for ADPG-PP from higher plants is being studied using a similar chemical labeling strategy. The chemical probe pyridoxal phosphate, which mimics the activator, 3-PGA, labels both the large (54 kD) and small (51 kD) subunits of the spinach leaf enzyme (Ball and Preiss, *J. Biol. Chem.* **269**:24706-24711, 1994; Morrell *et al.*, *J. Biol. Chem.* **263**:633-637, 1988). When pyridoxal phosphate is covalently bound, ADPG-PP no longer requires 3-PGA for maximal enzyme activity, suggesting that pyridoxal phosphate is bound at the activator site. Only a single reactive Lys, located near the C-terminus, is observed for the small subunit (Morell *et al.*, *J. Biol. Chem.* **263**:633-637, 1988). In contrast, three Lys residues are phosphopyridoxylated in the spinach leaf large subunit (Ball and Preiss, *J. Biol. Chem.* **269**:24706-24711, 1994). Two Lys residues are located near the C-terminus. One Lys is aligned with the conserved Lys labeled in the small subunit. A second reactive Lys residue is positioned 38 residues away from the former on the N-terminal side of the primary sequence. The third reactive Lys is located about 115-120 residues from the

N-terminus of the large subunit sequence and is not conserved in the small subunit sequence.

Using a random mutagenesis strategy, Green *et al.* (*Proc. Natl. Acad. Sci. USA* **93**:1509-1513, 1996) have shown that a mutation at Pro_{52} of the large subunit, when co-expressed with the wild-type small subunit, results in the formation of an ADPG-PP enzyme that is defective in allosteric regulation, requiring 45-fold greater levels of the activator 3-PGA than the wild-type enzyme. Pro_{52} is part of the sequence motif, PAV, that is conserved in all known ADPG-PPs (Smith-White and Preiss, *J. Mol. Evol.* **34**:449-464, 1992). Four residues away from the PAV in the *E. coli* enzyme is Lys_{39}, a residue that is located at or near the activator binding site in the bacterial enzyme. The small subunit has a Lys at the equivalent position of Lys_{39}, while the large subunit has an Arg, indicating a conservation of charge. The PAV motif also appears to be important for the allosteric response in the bacterial enzyme. When the Ala residue is replaced by a Thr, the resulting mutated enzyme has a greatly reduced affinity for the activator fructose-1,6-diphosphate. Overall, the results from chemical labeling and random mutagenesis studies indicate that sequences located at both the N- and C-terminal regions are required for allosteric regulatory behavior of higher plant ADPG-PP.

Roles of the large and small subunits in enzyme function. Under certain growth conditions, the small subunit alone can assemble to form a homotetrameric enzyme when expressed in *E. coli* (Ballicora *et al.*, *Plant Physiol.* **109**:245-251, 1995). The homotetrameric small subunit enzyme exhibits values of K_m for glucose 1-phosphate, ATP and Mg^{2+} ions that are similar to those of the wild-type heterotetrameric enzyme. The homotetrameric small subunit enzyme, however, requires at least 15-fold greater amounts of 3-PGA for 50%

activation than the wild-type enzyme. In contrast, the large subunit alone is unable to assemble into a catalytically active enzyme. These observations suggest that the two subunit types possess distinct roles in the functioning of this enzyme. The small subunit may play a more dominant role in catalysis, while the large subunit may increase the sensitivity of the small subunit to activation by 3-PGA. Results from random mutagenesis (Greene *et al.*, *Proc. Natl. Acad. Sci. USA* **93**:1509-1513, 1996; Greene *et al.*, *Plant Physiol.* **112**:1315-1320, 1996; Laughlin *et al.*, *Phytochem.* **47**:621-629, 1998) of the large and small subunit sequences support this view. Most large subunit mutants that display high enzymatic activity are defective in allosteric regulation, while high-activity small subunit mutants are mainly defective in binding of the substrates glucose 1-phosphate and/or ATP.

ADPG-PPs of Seeds and Tubers. In addition to the genes that encode the leaf ADPG-PP, cereals and peas possess a second set of genes that encode a seed-specific form. For example, the maize endosperm-specific ADPG-PP is composed of two distinct subunits encoded by the *Bt2* and *Sh2* loci. Immunoblot analysis of protein extracts revealed that the bt_2 mutant lacked a 55-kD species, while the sh_2 mutant lacked a 60-kD species (Preiss *et al.*, *Plant Physiol.* **92**:881-885, 1990). The identification of the maize ADPG-PP subunits was substantiated by structural analysis of *Sh2* and *Bt2* genes, which revealed substantial homology to known ADPG-PP sequences from both plants and bacteria (Bae *et al.*, *Maydica* **35**:317-322, 1990; Bhave *et al.*, *Plant Cell* **2**:581-588, 1990; Smith-White and Preiss, *J. Mol. Evol.* **34**:449-464, 1992). Based on these results, *Bt2* and *Sh2* contain the structural gene sequences for the maize

endosperm ADPG-PP small and large subunits, respectively.

In contrast to the absolute dependence of the leaf ADPG-PP for the activator 3-PGA, ADPG-PPs from storage tissues display variable allosteric responses *in vitro*. ADPG-PPs from maize endosperm (Plaxton and Preiss, *Plant Physiol.* **83**:105-112, 1987) and potato tubers (Sowokinos and Preiss, *Plant Physiol.* **69**:1459-1466, 1982) are absolutely dependent on 3-PGA for maximum enzyme activity. Substantial enzyme activity is evident only when the enzyme is assayed in the presence of 3-PGA, yielding more than a 25-fold activation over levels observed in the absence of activator. Likewise, the rice endosperm enzyme displays allosteric activation by 3-PGA, although the level of activation (5-fold) is much less than the maize enzyme. In contrast, ADPG-PPs from barley (Kleczkowski *et al.*, *Plant Physiol.* **101**:179-186, 1993), wheat endosperm (Duffus, *Biochem. Soc. Trans.* **20**:13-18, 1992; Olive *et al.*, *Plant Mol. Biol.* **12**:525-538, 1989), and pea embryos (Hylton and Smith, *Plant Physiol.* **99**:1626-1634, 1992) display little or no allosteric regulation. In each instance, substantial enzyme activity was evident even in the absence of 3-PGA, and the enzyme was not significantly activated by 3-PGA or inhibited by Pi, suggesting that ADPG-PPs from these developing seeds may not be subject to allosteric control, unlike leaf ADPG-PPs.

Alternatively, the lack of allosteric response exhibited by the barley, wheat and pea seed enzymes may not be an intrinsic property of these enzymes but rather a result of a post-translational modification of the enzyme. Post-translational-induced changes in allosteric properties of these ADPG-PPs is a likely possibility, as suggested by studies of the maize endosperm enzyme. When maize endosperm extracts are prepared in the absence of proteinase inhibitors, a

substantial level of enzyme activity is observed even in the absence of the activator 3-PGA. In addition, the sensitivity to allosteric effectors was observed to be much less than the leaf form. Only a two- to three-fold activation is detected under optimal conditions as opposed to the 20- to 30-fold evident for the leaf enzyme (Dickinson and Preiss, *Arch. Biochem. Biophys.* **130**:119-128, 1981). ADPG-PP enzyme from maize endosperm has been found to be highly susceptible to proteolysis (Plaxton and Preiss, *Plant Physiol.* **83**:105-112, 1987). Incubation of crude extracts at 30°C resulted in the degradation of the Bt2 (55 kD) subunit to a 53 kD entity as viewed by immunoblot analysis using antibody raised against the spinach leaf enzyme. When the enzyme was purified in the presence of protease inhibitors, it was shown to have allosteric and physical properties similar to those of leaf ADPG-PPs. In contrast to the proteolytically cleaved enzyme which displays very little allosteric response, the intact enzyme was activated by about 25-fold by 3-PGA and this activation was suppressed by Pi.

An important conclusion drawn from the study of the maize endosperm enzyme (Plaxton and Preiss, *Plant Physiol.* **83**:105-112, 1987) is that small structural changes of the native enzyme result in significant changes in the catalytic and allosteric properties of enzyme function. Such proteolytic-induced changes are also evident for the recombinant potato enzyme expressed in *E. Coli* (Iglesias *et al.*, *J. Biol. Chem.* **268**:1081-1086, 1993) and are believed to be responsible for the distinct properties exhibited by the barley and wheat endosperm enzyme activities as well as those from pea embryos (Preiss *et al.*, "Prospects for the production of cereals with improved starch properties," In: *Improvement of Cereal Quality by Genetic Engineering*, Henry and Ronalds, eds., Plenum Press, 1994).

Expression of a recombinant potato tuber enzyme whose small subunit is lacking about nine residues of the N-terminus results in an enzyme that is less stable to heat denaturation and is only sensitive to Pi inhibition in the presence of 3-PGA. Restoration of these nine residues on the small subunit results in the formation of a recombinant enzyme that is stable toward heat treatment at 60°C and more sensitive to Pi inhibition. As mentioned above, the barley endosperm enzyme displays almost no activation by 3-PGA and is only weakly inhibited by Pi (Kleczkowski *et al.*, *Plant Physiol.* **101**:179-186, 1993). Immunoblot analysis of crude extracts indicated, however, that even at 4°C, the barley large subunit was degraded first to a 53 kD and then later to a 51 kD species with a half-life on the order of minutes. In view of the distinct changes in allosteric and catalytic properties mediated by the structural changes of the maize endosperm and the recombinant tuber enzymes, the distinct allosteric properties exhibited by the barley endosperm enzyme (and wheat and pea as well) are likely due to proteolysis of one or both subunits (Preiss *et al.*, "Prospects for the production of cereals with improved starch properties," In: *Proceedings of th Improvement of Cereal Quality by Genetic Engineering*, Henry and Ronalds, eds., Plenum Press, 1994).

Irrespective of the *in vitro* allosteric properties displayed by the seed and tuber enzymes, questions have been raised regarding the functional significance of this control mechanism *in vivo*. Unlike leaf tissue, where starch metabolism must be tightly regulated so that synthesis occurs during the day and is degraded at night, synthesis and degradation of starch in seeds and tubers are temporally separated to distinct stages of plant development. Therefore, allosteric control of ADPG-PP during the diurnal cycle is not an essential requirement for starch synthesis in non-photosynthetic

sink tissues. Indeed, the inherent allosteric property of ADPG-PP may obstruct starch synthesis as the levels of the activator 3-PGA are likely to be much lower than the inhibitor Pi in amyloplasts due to the unique biochemistry of this organelle. Unlike the more autonomous chloroplast, which generates its own carbon and energy via CO_2 fixation and photophosphorylation (which result in a production of 3-PGA and ATP and a corresponding reduction in Pi), the amyloplast is dependent on the cytoplasm for these metabolic requirements.

These biochemical differences are also reflected in the permeability properties of chloroplasts and amyloplasts. In chloroplasts, the principal metabolite transport system between the chloroplast and cytoplasm is the Pi translocator, which transports triose phosphate and 3-PGA to the cytoplasm in exchange for Pi. In contrast, in non-photosynthetic sink organs, carbon is imported in the form of sucrose which is subsequently processed into hexose-phosphate, which is transported into the amyloplast to be directly utilized as a substrate (Glc-1-P) by ADPG-PP (Heldt *et al.*, *Plant Physiol.* **95**:341-343, 1991; Hill and Smith, *Planta* **185**:91-96, 1991; Keeling *et al.*, *Plant Physiol.* **87**:311-319, 1988; Kosegarten and Mengel, *Physiol. Plant.* **91**:111-120, 1994; Okita, *Plant Physiol.* **100**:560-564, 1992; Tyson and ap Rees, *Planta* **175**:33-38, 1988) or is converted into ADPglucose in the cytoplasm by a cytoplasmically-localized ADPG-PP (Denyer et al., *Plant Physiol.* 112: 779-785, 1996; Thorbjonsen et al., Plant J. 10: 243-250, 1996). The ADPglucose is then transported to the plastid where it is then utilized by starch synthase.

These proposed pathways and the growing evidence that amyloplasts lack intact glycolytic and gluconeogenic pathways (Entwistle and ap Rees, *Biochem.*

J. **271**:467-472, 1990; Frehner *et al.*, *Plant Physiol.* **94**:538-544, 1990) indicate a limited role for triose-phosphate in amyloplastic starch biosynthesis. Although 3-PGA is probably not a major metabolite in the amyloplasts, Pi is likely a major metabolite because it is a direct byproduct of the ADPG-PP reaction (in conjunction with inorganic pyrophosphatase), resulting in low 3-PGA/Pi ratios. Under these conditions, the allosteric properties of the maize and rice endosperm and tuber ADPG-PP enzymes suppress the catalytic activities of these enzymes and, in turn, constrain starch synthesis in these plants.

Evidence in support of this view has been obtained in tubers. First, tuber disks incorporate ^{14}C-labeled sucrose into starch at 50% higher rates in the presence of mannose, which sequesters the inhibitor Pi (Hnilo and Okita, *Plant and Cell Physiol.* **30**:1007-1010, 1989). Second, expression of an allosteric *E. coli* mutant of ADPG-PP results in an enhancement of starch synthesis in tubers (Stark *et al.*, *Science* **258**:287-292, 1992). Third, a maize line, *Sh2-m1Rev6*, which bears seeds with weight increases of 10-18% higher than normal lines, encodes a ADPG-PP large subunit that contains two additional amino acids near the C-terminus (Giroux *et al.*, *Proc. Natl. Acad. Sci. USA* **93**:5788-5792, 1996), a peptide region known to be essential for allosteric regulation. The resulting variant ADPG-PP enzyme appears to be resistant to Pi inhibition, suggesting that the altered allosteric response may be responsible for increase seed weight. Overall, these results support the view that starch synthesis has not reached it highest potential in developing harvestable sink organs, i.e. tubers, seeds and fruit, and that expression of allosteric mutant plant ADPG-PPs would increase starch synthesis and, in turn, plant yields.

It would be desirable and advantageous to produce a mutant form of the ADPG-PP enzyme wherein a plant expressing such a gene would have increased yield and increased plant size.

In addition to having a large impact on the degree of starch synthesis and, in turn, yield in developing storage organs such as developing seeds, tubers, fruit etc., ADPG-PP activity may also have an important role in increasing overall productivity of the plant by maximizing rates of carbon dioxide (CO_2) fixation and utilization in photosynthetic tissues such as leaves, stems etc. In general, CO_2 fixation rates in photosynthetic tissues is affected by the capacity of converting fixed CO_2 into carbohydrates such as sucrose and starch. Recent evidence indicates that leaf starch plays a broader role than simply serving as a diurnal reserve of carbon and energy to enable the plant to survive during the dark period. Ludewig et al. (*FEBS Lett.* **429**:147-151) have shown that there is a direct correlation between the capacity of starch synthesis and the rate of photosynthesis at elevated CO_2. Likewise, recent evidence from J. Sun, G.E. Edwards, and T.W. Okita (unpublished) showed that there are significant correlations between the rates of starch synthesis and CO_2 assimilation, and between the rates of starch synthesis and accumulative leaf area. These results indicate that leaf starch plays an important role as a transient "sink" in which synthesis can ameliorate potential reduction in photosynthesis due to feedback regulation.

SUMMARY OF THE INVENTION

Applicants have discovered methods for generating and identifying up-regulated mutants of ADPG-PP, i.e., allosterically regulated mutant ADPG-PPs having enzymatic activity that is significantly higher than an otherwise similar wild-type enzyme at physiological

concentrations (i.e., about equimolar levels) of 3-PGA and Pi (Heldt *et al.*, *Plant Physiol.* **59**:1146-1155, 1977). The methods of the invention are more efficient than any previous empirical methods involving mutagenesis and mass biochemical screening of mutants.

Applicants have generated such ADPG-PP mutants and have analyzed the enzymatic activity of the mutant ADPG-PP enzyme. Applicants have discovered that, in comparison to the wild-type enzyme, the mutant enzyme differs in that it has: (1) higher sensitivity to the activator 3-PGA; (2) lower sensitivity to the inhibitor Pi; (3) increased starch production; (4) increased yield; (5) increased plant size (especially leaves); and, surprisingly, (6) increased growth rate and (7) increased number of seeds.

According to one aspect of the invention methods are provided for producing and identifying nucleic acids that encode up-regulated mutant ADPG-PP enzymes. For example, in a first representative embodiment of the method, the plant ADPG-PP large and small subunits are co-expressed in a bacterial cell lacking ADPG-PP activity (for example a $glgC^-$ strain). One subunit is unmutated while the other is mutated. If the mutation affects only the allosteric regulatory properties of the assembled enzyme, then cells expressing this mutant enzyme will not accumulate glycogen under normal physiological levels of the activator 3-PGA, although they will have normal levels of ADPG-PP activity if measured in the presence of saturating levels of 3-PGA. A simple means for assessing glycogen levels, which can be correlated with ADPG-PP anzyme activity levels in a bacterial cell, is by exposing the cells to iodine vapor. Glycogen-accumulating cells stain brown to purplish black while cells lacking this carbohydrate stain light yellow. Upon determining that the mutated subunit affects only the allosteric regulatory properties of the assembled enzyme, the mutated subunit is again mutagenized to produce a second mutation which

will restore the allosteric regulatory properties of the assembled enzyme and the bacterial cell's capacity to accumulate glycogen which can be readily detected by iodine staining. A subunit that contains only the second mutation but not the first mutation is then studied to determine the effect of this second mutation on ADPG-PP function, for example, whether the second mutation is an up-regulatory mutation.

An alternative embodiment of a method for the generation and identification of up-regulated ADPG-PP mutants is by direct screening methods. However, in order to directly detect up-regulated mutants, the plant ADPG-PP subunits must be expressed on low-copy-number plasmids. Alternatively, high-copy-number plasmids can be used to co-express the plant ADPG-PP subunits; but, the bacterial cells must be grown on enriched media (Govons et al., *J. Bacteriol.* **97**:970-972, 1969) containing low amounts of glucose (0 to 0.5%) as compared to the normal 2%. When grown on enriched media containing 0.1% glucose, cells expressing up-regulated ADPG-PP readily stain with iodine vapor whereas cells expressing the wild-type unmutated enzyme do not.

A third representative method for generating and identifying up-regulated ADPG-PP mutants involves generation of segmented mutations. Nucleotides encoding ADPG-PP are truncated such that the expressed proteins are truncated at either the N- or C- terminus. These mutants are screened for altered allosteric function and up-regulated mutants are identified by methods described herein.

According to another aspect of the invention, nucleic acids are provided that encode an up-regulated mutant ADPG-PP enzyme, for example, a plant mutant ADPG-PP enzyme. Such a nucleic acid may be made and identified by any of the methods of the invention. Alternatively, once a particular mutant sequence has been identified, the nucleic acid may be made by chemical *in vitro* synthesis, such as by using an

automated polynucleotide synthesizer, or may be made using site-directed mutagenesis. The mutant ADPG-PP nucleic acid codes for an enzyme that is up-regulated, and therefore has a biological activity (i.e., an enzyme activity) that is higher than the biological activity of the wild-type enzyme under physiological conditions.

According to yet another aspect of the invention, enzymes are provided that are encoded by the nucleic acids of the invention, for example, a plant mutant ADPG-PP enzyme. Such enzymes are up-regulated mutant enzymes. For instance, such a mutant enzyme may be more sensitive to an allosteric activator, such as 3-PGA, and may be more resistant (i.e., less sensitive) to an allosteric inhibitor such as Pi. Up-regulated mutant enzymes may be produced by a number of methods including, but not limited to, mutagenesis with chemicals such as hydrozylamine, mutagenesis with radiation, site-directed mutagenesis, PCR mutagenesis, or genetic engineering techniques such as N-terminal truncation of the enzyme.

According to yet another aspect of the invention transgenic plants are provided that comprise the mutant nucleic acids of the invention. Expression of up-regulated ADPG-PP mutant genes in transgenic plants can result in, for example, increased starch production in at least a portion of the plant (compared with the wild-type) and can also increase productivity, yield, growth rate, and seed number.

According to yet another aspect of the invention specific nucleic acids as recited in the specification are provided.

The foregoing and other aspects of the invention will become more apparent from the following detailed description and accompanying drawings.

SEQUENCE LISTING

SEQ ID NO:1 shows the nucleotide sequence of the mutant ADPG-PP large subunit UpReg1 that, when co-

expressed with a wild-type ADPG-PP small subunit, forms an up-regulated ADPG-PP enzyme.

SEQ ID NO:2 shows the primary amino acid sequence of UpReg1.

SEQ ID NO:3 shows the nucleotide sequence of the mutant ADPG-PP large subunit UpReg2, that when co-expressed with a wild-type ADPG-PP small subunit, forms an up-regulated ADPG-PP enzyme.

SEQ ID NO:4 shows the primary amino acid sequence of UpReg2.

SEQ ID NO:5 shows the nucleotide and amino acid sequences of ΔN17-LS, a mutant ADPG-PP large subunit that, when co-expressed with a wild-type ADPG-PP small subunit, forms an up-regulated ADPG-PP enzyme.

SEQ ID NO:6 shows the primary amino acid sequence of ΔN17-LS.

SEQ ID NO:7 shows the nucleotide sequence of the mutant ADPG-PP large subunit R20 that, when co-expressed with a wild-type ADPG-PP small subunit, forms an up-regulated ADPG-PP enzyme.

SEQ ID NO:8 shows the primary amino acid sequence of R20.

SEQ ID NO:9 shows the nucleotide sequence of the mutant ADPG-PP large subunit R32 that, when co-expressed with a wild-type ADPG-PP small subunit, forms an up-regulated ADPG-PP enzyme.

SEQ ID NO:10 shows the primary amino acid sequence of R32.

SEQ ID NO:11 shows the nucleotide sequence of the mutant ADPG-PP large subunit R4 that, when co-expressed with a wild-type ADPG-PP small subunit, forms an up-regulated ADPG-PP enzyme.

SEQ ID NO:12 shows the primary amino acid sequence of R4.

SEQ ID NO:13 shows the nucleotide sequence for an upstream primer 5'-GATATTGGTACCATTG-3' that is useful for introducing double-termination codons in an ADPG-PP

small subunit cDNA sequence. The primer includes a KpnI site.

SEQ ID NO:14 shows the nucleotide sequence for a downstream primer that is useful for introducing double termination codons in an ADPG-PP small subunit cDNA sequence. The primer includes a SacI site. Termination codon sequences are underlined.

SEQ ID NO:15 shows the nucleotide sequence for an upstream primer that is useful for introducing double termination codons in an ADPG-PP large subunit cDNA sequence. The primer includes an NheI site.

SEQ ID NO:16 shows the nucleotide sequence for an downtream primer that is useful for introducing double-termination codons in an ADPG-PP large subunit cDNA sequence. The primer includes a SacI site. Termination codon sequences are underlined.

SEQ ID NO:17 shows the nucleotide sequence for an upstream primer that is useful for restoring the N-terminus of an ADPG-PP small subunit expression plasmid. The primer includes an NcoI site.

SEQ ID NO:18 shows the nucleotide sequence for a downstream primer that is useful for restoring the N-terminus of an ADPG-PP small subunit expression plasmid. The primer includes a KpnI site.

SEQ ID NO:19 shows the nucleotide sequence for an upstream primer that is useful for deleting DNA sequences that code for 17 amino acids at the N-terminus of an ADPG-PP large subunit. The primer includes an NcoI site.

SEQ ID NO:20 shows the nucleotide sequence for a downstream primer that is useful for deleting DNA sequences that code for 17 amino acids at the N-terminus of an ADPG-PP large subunit. The primer includes an NheI site.

SEQ ID NO:21 shows the upstream primer used to amplify the *Arabidopsis* ribulose bisphosphate small subunit (ats1A) promoter and transit leader coding sequences.

SEQ ID NO:22 shows the downstream primer used to amplify the *Arabidopsis* ribulose bisphosphate small subunit (ats1A) promoter and transit leader coding sequences.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 shows the structure of the plasmid pML7.

FIG. 2 shows the structure of the plasmid pML10.

FIG. 3 shows a comparison of the I_2 staining patterns of AC70R1-504 cells expressing the wild-type (LS/SS) and the mutant UpReg-1, R20, and 345 ADPG-PP enzymes when grown on glucose-enriched media. Mutant 345, which contains the P52L mutation on the large subunit (LS) does not accumulate glycogen and stains a faint yellow. In contrast, wild-type (LS/SS) ADPG-PP and the mutants UpReg-1 and R20 accumulate large amounts of glycogen and stain darkly. Note the more intense staining by UpReg-1 cells as compared to the wild-type (LS/SS) and R20 cells.

FIG. 4 shows a graphical comparison of $I_{0.5}$ (the amount of Pi required for 50% inhibition) values at different 3-PGA concentrations for the native, recombinant wild-type (LS+SS), and ΔN17-LS ADPG-PP enzymes. Data on the native enzyme were taken from Sowokinos and Preiss, *Plant Physiol.* **69**:1459-1466, 1982, while data for the recombinant enzyme were taken from Ballicora *et al.* (*Plant Physiol.* **109**:245-251, 1995, and from Laughlin and Okita (*Phytochem* **47**:621-629, 1998).

FIG. 5 is a table comparing substrate binding (Km) 3-PGA activator ($A_{0.5}$) affinity properties of various native and mutant ADPG-PPs.

FIG. 6 is a table comparing Pi inhibition ($I_{0.5}$) of various native and mutant ADPG-PPs.

FIG. 7 is a table comparing 3-PGA activation for various native and mutant ADPG-PPs. $A_{0.5}$ is the amount of 3-PGA required to give 50% activation. $I_{0.5}$ is the

amount of Pi required to inhibit the enzyme 50% in the presence of a known amount of 3-PGA.

FIG. 8 is a table showing seed yields from eight greenhouse-grown plants of various plant lines (T3 generation).

Fig. 9 shows a recombinant vector comprising the ats1A-potato large subunit cassettes contained within a Xba I/Sac I DNA fragment cloned into the Xba I and Sac I sites of the T-DNA binary vector pHI-32, a derivative of pIG-121, to produce pHI-33 to pHI-39.

Fig. 10 shows third generation progeny of the transgenic *Arabidopsis* plants containing pHI-33 to pHI-39 germinated and cultured on MS media. Transgenic plants expressing up-regulatory AGPase (left and right) grew considerably faster than control plants (center).

Fig. 11 shows a generalized schematic drawing of the plasmid used to construct a gene fusion between the ats1A promoter and transit leader sequences to the potato ADPG-PP large subunit sequences. The coding sequences from UpReg1, R4, R20, R32, M27, M345 and wildtype large subunit sequences were removed from the plasmid DNA by digestion with Nco I and Sac I, and the resulting DNA fragment was cloned into the relevant restriction sites of pHI-10 to produce plasmids pHI-11 to pHI-17.

Fig. 12 shows the *Arabidopsis* ribulose bisphosphate small subunit ast1A promoter and transit leader coding sequence fragment digested with Xho I and Sac I cloned into the Xho I and Sac I sites of pBluescript II.

Fig. 13 is a table comparing phenotypic properties of various transgenic plants.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Overall yield or productivity is governed not only by a plant's capacity to fix carbon dioxide and produce energy but also by the ability of the plant to utilize this fixed carbon efficiently. For many crop plants

that accumulate starch in their reserve organs, productivity is limited by the capacity of the plant to convert sugar into starch. ADPG-PP, which controls the flux of carbon into starch, is allosterically regulated.

The level of ADPG-PP activity is controlled by the levels of the effectors 3-PGA and Pi, which activate and inhibit, respectively, the activity of ADPG-PP. Because there are significant levels of Pi in the cell and levels of this effector molecule are not expected to diurnally oscillate as in photosynthetic tissue, the net enzyme catalytic activity and, in turn, the net rate of starch synthesis in developing sink organs such as tubers and developing seeds, is much lower than the maximum potential rate.

In addition to the above-summarized biochemical events in developing non-photosynthetic sink organs such as tubers, developing seeds, and fruits, etc., starch synthesis in photosynthetic tissues such as leaves may also impact productivity. Recent evidence by Ludewig et al. (*FEBS Lett.* **429**:147-151) and unpublished results (Jindong Sun, G.E. Edward and T.W. Okita) indicate that there is a direct correlation between starch synthesis and CO_2 fixation. These observations indicate that starch can serve as a transient "sink" to store fixed carbon in a chemical form that does not inhibit CO_2 fixation and photosynthesis. Because ADPG-PP is a key regulatory enzyme in controlling starch synthesis, allosteric regulatory mutant forms of this enzyme may increase leaf starch formation and, in turn, productivity.

Three approaches have now been discovered for generating and identifying variant ADPG-PPs that display up-regulated allosteric or different physical properties, e.g., temperature resistance. These allosterically up-regulated mutants require substantially less 3-PGA to attain maximum catalytic activity and/or are less susceptible to Pi inhibition than the unmutated enzyme. Mutant ADPG-PP genes are

useful, for example, for altering starch production in a host organism, e.g., to increase starch production in cereal seeds, tubers, fruit, leaves and other plant organs and tissues. Yield is also increased. Surprisingly, certain mutant ADPG-PP genes, when expressed in a plant, result in increased growth rate and increased seed production.

A first approach involved the following steps:

(1) A native ADPG-PP gene was mutagenized to produce a single mutation.

(2) Mutants that are defective in enzyme function were identified by their inability to complement a mutation in the bacterial *glgC* gene. Cells harboring these mutant sequences do not produce glycogen and hence are not stained with iodine vapor.

(3) The mutant cells were cultured and the mutant ADPG-PP enzyme was purified. Enzyme kinetics were determined by standard methods (Cornish-Bowden, *Analysis of Enzyme Kinetic Data*, Oxford University Press, 1995); thus, enzymes were identified that are defective in allosteric function (Green et al., 1996).

(4) Having been identified, the allosteric mutants were then subjected to a second round of mutagenesis to produce a double mutant.

(5) Double mutant cells were then screened for staining by iodine vapor, i.e. a reversal of the iodine staining-minus phenotype that is mediated by the primary mutation.

(6) Double mutants exhibiting at least partial complementation of *glgC*$^-$ cells were then sequenced to reveal the location and nature of the two mutations.

(7) Site-directed mutagenesis was then used to create a DNA containing the second (but not the first) mutation.

(8) The protein product of the DNA containing the secondary site mutation was then analyzed by standard enzyme-kinetics methods (Cornish-Bowden, 1995) in the

absence of the initial primary mutation to determine whether the secondary site mutation modifies the allosteric properties of the enzyme by itself. Mutants were thus identified that had increased biological activity over the wild-type enzyme and were therefore identified as "up-regulated" mutant enzymes.

A second approach involved the direct screening method to identify up-regulated allosteric mutants. Cells containing an up-regulated mutant enzyme would be expected to have increased glycogen production. However, using the iodine vapor staining method to assess mutations, up-regulated mutants from normal cells are not distinguishable because the iodine screening procedure is saturated. That is, although cells harboring the up-regulated mutant accumulate more glycogen than cells containing the wild-type enzyme, the up-regulated cells stain with iodine at the same rate as cells containing the wild-type enzyme. To circumvent this barrier, the ADPG-PP sequences were placed on a low copy-number plasmid. The plamids used were pWSK28 (that contains the polylinker from pBluescript Skt) and pWSK30 (that contains the polylinker from pBluescript Kst).

Low copy-number plasmids are discussed in standard texts (e.g., Sambrook et al., 1989) and may be obtained commercially. Low copy-number plasmids other than pWSK28 and pWSK30 include pBI101 (Jefferson et al., *EMBO* 6:3901-3907, 1987) which has a wide host-range replicon.

Under normal growth conditions, the level of enzyme produced is insufficient to completely complement the *glgC* mutation and the cells stain only lightly with iodine vapor. On the other hand, cells containing an up-regulated mutant enzyme are more catalytically active and greater glycogen levels will be produced. As a result, cells with an up-regulated mutant enzyme stain darker with iodine.

Alternatively, instead of changing the copy number of the expression plasmids, cells expressing up-

regulatory mutant ADPG-PP can be distinguished from cells containing the wild-type enzyme by altering the glucose concentration in the enriched media (Govons et al., *J. Bacteriol.* **97**:970-972, 1969). When cultured in enriched media containing 0.1% glucose instead of the usual 2% glucose level, up-regulatory mutants will stain darkly with I_2, whereas normal cells stain very lightly.

A third approach involved the generation of allosteric mutants by segmental mutations. DNA corresponding to N- and or C-terminal portions of the ADPG-PP polypeptide were deleted by recombinant DNA techniques (Sambrook et al., 1989). Deletions may include the removal of nucleotides corresponding to 5, 10, 15, 17, 25, 30, 40 or more amino acid residues from the N- or C- terminus.

In one embodiment of the invention, *Escherichia coli* expression systems were used to express the large and small subunit cDNAs of ADPG-PP under the control of bacterial promoters. For example, the large subunit cDNA was cloned into a pACYC-based cloning vector with expression of the large subunit cDNA being driven by a *tac* promoter, while the small subunit cDNA was cloned in a pBR325-based cloning vector with expression of the small subunit cDNA driven by a *recA* promoter. Examples of plasmid vectors useful for the expression of the large and small subunit cDNAs can be found, for example, in Iglesias *et al.* (*J. Biol. Chem.* **268**:1081-1086, 1993).

Many suitable expression systems are commercially available, from, for instance, Invitrogen, Pharmacia and New England Biolabs.

Those skilled in the art will recognize that co-expression of the large and small subunit cDNAs can be accomplished with a wide variety of compatible hosts and cloning vectors and that both the large and small subunit genes can be cloned and expressed together on a single vector.

The Examples below discuss up-regulated mutant ADPG-PP enzymes that require less than one-tenth of the level of 3-PGA than wild-type ADPG-PP for complete activation. These mutant enzymes enhance the rate of glycogen synthesis in bacteria and can increase starch synthesis in other organisms such as plants if expressed in the appropriate tissue. Expression systems that can be used to express genes in plants are widely known and commercially available. For instance, the *Agrobacterium* system may be used to carry out transformation using pHI-32 or pCAMBIA T-DNA plasmids, both of which are derived from pBI101. Such plasmids can be used to transform plants such as potato, rice, wheat, barley, maize, and tomato.

The mutant ADPG-PP subunits generated by this invention can be used to increase yield and productivity of many crop plants that use starch as their principal reserve in their photosynthetic tissues and/or storage organs. Expression of these mutant subunits and subsequent formation of ADPG-PP enzymes with up-regulatory properties can be used to increase starch production in storage organs such as potato tubers, cassava roots, and cereal seeds including those from wheat, maize, rice, barley, rye, and sorghum. The mutant subunits can be used to increase starch production during the early development of the storage organs in oil-accumulating seed plants such as soybean, rape, and sunflower and in developing fruits of tomato, apple, pear, peach, etc. In both oil-accumulating seed plants and in fruiting plants, starch is used as a transient reserve of carbon and energy which is then re-utilized for the formation of other molecules. In oil-accumulating seed plants, starch accumulates during the early phase of seed development and is then re-utilized for the production of oils. In fruiting plants, accumulated starch is metabolized to reducing sugars, a preferred trait for tomatoes especially those used for

processing ketchup, paste and the like and for the sweetness of apples, pears and other sweet fruits.

In addition to the engineering of storage organs, mutant ADPG-PP subunits generated by this invention can be used to increase overall productivity of all C_3 plants (defined as those plants that use the Calvin cycle to fixed CO_2 into 3-PGA as their first stable three-carbon intermediate). Examples of C_3 plants include potato, cassava, wheat, rice, barley, rye, soybean, rape, sunflower, flax, cotton, alfalfa, celery, cauliflower, and carrot. C_3 plants typically exhibit CO_2-limited photosynthesis at ambient and sub-ambient CO_2 partial pressures. Under ideal conditions, the rate of C_3 photosynthesis increases with increasing CO_2 as photorespiration is suppressed. However, under many environment conditions, e.g. low temperatures, high light, and increasing CO_2 (see Leegood and Edwards, *Photosynthesis and the Environment*, N.R. Baker, ed., Kluwer Academic Publ., pp. 191-221, 1996) the rate of photosynthesis is much lower than predicted because of the limitations in triose phosphate utilization. An excellent example that illustrates this "feedback" of photosynthesis is rice. Rice grown at a photosynthetic photon flux of 1000 μM m^{-2} s^{-1} and 26°C day, 24°C night exhibits rates of photosynthesis significantly lower than expected when shifted to low levels of O_2 under moderate temperatures and ambient levels of CO_2 (Sun, Edwards and Okita, unpublished data), a response indicative of photosynthetic feedback. The limitations in triose phosphate utilization can be overcome if starch synthesis is elevated in photosynthetic-competent tissues such as leaves and photosynthetic-competent tissues such as stems and seed pods, or grains. In turn, this increased photosynthesis will result in higher biomass production (productivity) and, in turn, higher yields of harvestable organs such as tuber, seeds, fruit etc.

The mutant ADPG-PP subunits can be transferred into crop plants by either *Agrobacterium*-mediated transformation or by mechanical introduction, e.g., biolistics bombardment. *Agrobacterium*-mediated transformation is the preferred method because of the smaller frequency of introducing multiple copies of the transgene and subsequent problems of gene expression instability due to co-suppression effects in offspring of transgenic plants (Kumpatla et al., *Plant Phsyiol.* **115**:361-373, 1997). Agrobacterium has been used to introduce transgenes into rice (Hiei et al.,*Plant J.* **6**:271-282, 1994), maize (Ishida et al. *Nature Biotechnol.* **14**:745-750, 1996; Gould, *Plant Physiol.* **95**:426-434, 1991), soybean (Hinchee et al. *Biotechnology* **6**:915-922, 1988; Stewart et al., *Plant Physiol.* **112**:121-129, 1996), rape (Falco et al., *Biotechnology* **13**:577-582, 1995), potato (Stark et al., *Science* **258**:287-292, 1992; Tu et al. *Plant Mol. Biol.* **37**:829-838, 1998), tomato (McGarvey et al., *Biotechnology* **13**:1484-1487, 1995), sunflower (Bidney et al., *Plant Mol. Biol.* **18**:301-313, 1992), cotton (Hansen et al., *Proc. Natl. Acad. Sci.* USA **91**:7603-7607, 1994), cassava (Li et al., *Nature Biotechnol.* **14**:736-740, 1996), apple (Maximova et al., *Plant Mol. Biol.* **37**:549-559, 1998). Pea (Lurquin et al., *Mol. Biotechnol.* **9**:175-179, 1998), and chickpea (Ramana et al., Indian *J. Exp. Biol.* **34**:716-718, 1996) as well as lettuce, sugarbeet, celery, cucumber, alfalfa, carrot, cauliflower, horseradish, poplar, walnut and asparagus (Gasser and Fraley, *Science* **244**:1293-1299, 1989). Wheat and barley have also been transformed by **Agrobacterium** (D. von Wettstein, personal communication).

An important consideration for the engineering of these plants with the mutant ADPG-PP subunits is the judicious use of promoters to drive expression of the introduced transgene at high levels in targeted tissues. Examples include the use of the patatin promoter to drive expression in potato tubers (Stark et al., *Science* **258**:287-292, 1992), the vicilin or kunitz trypsin inhibitor promoter to drive expression during the early stages of seed development of soybeans and other legumes (Walling et al., *Proc. Natl. Acad. Sci.* USA **83**:2123-2127, 1986; Wandelt et al., *Plant J.* **2**:181-192, 1992) as well as chickpea, pea, and rape, the glutelin Gt1 (Okita et al., *J. Biol. Chem.* **264**:12573-12581, 1989), the wheat high molecular weight glutenin and its related *hor* D gene in barley (Blechl and Anderson, *Nature Biotechnol.* **14**:875-879, 1996), or Shrunken 2 (Shaw and Hannah, *Plant Physiol.* **69**: 1214-1219, 1992) promoters to drive expression in developing cereal seeds, the *ats1A* promoter to drive expression in leaves and other photosynthetic competent tissues (Krebbers et al., *Plant Mol. Biol.* **11**:745-759, 1988), the E4 or E8 promoters to drive expression in developing tomato fruit (Deikman et al., *Plant Cell* **1**:1025-1034, 1989), and the AX92 promoter to drive expression in root cortex (Dietrich et al., *Plant Cell* **4**:1371-1382, 1992).

Comparison with previous methods for manipulating ADPG-PP activity. Two previous approaches to manipulating ADPG-PP activity have been utilized. One approach (referred to herein as the "Monsanto approach"), utilizes a mutant bacterial ADPG-PP, glgC16, that codes for a form of ADPG-PP that lacks allosteric regulation. The glgC16 enzyme possesses 60% of the activity levels of the fully activated wild-type enzyme in the absence of the activator fructose-1,6-diphosphate and is also less sensitive to the inhibitor adenosine

monophosphate (AMP). Transfer of this gene and expression in tubers and tomatoes results in potatoes with higher starch content (*see* Stark *et al.*, *Science* **258**:287-292, 1992) and tomatoes with higher solids content.

A second approach utilizes the transposable element system Ac-Ds in maize to generate ADPG-PP variants. One mutant was obtained by excision of a Ds element that was initially located in the *Shrunken-2* gene, which encodes the endosperm-specific form of the large subunit gene of ADPG-PP. During excision of Ds, a six-nucleotide "footprint" was left behind, resulting in a large subunit having two additional amino acids. The resulting enzyme appears to be resistant to Pi inhibition. A modified potato large subunit containing these two additional residues formed an enzyme resistant to Pi inhibition (Giroux *et al.*, *Proc. Natl. Acad. Sci. USA* **93**:5824-5829, 1996).

The present invention differs from the Monsanto approach in that it utilizes (1) isolated plant genes for ADPG-PP instead of a bacterial gene and (2) an enzyme that is subject to allosteric regulation, although its sensitivity to activation by 3-PGA and inhibition by Pi is 20- to 40-fold higher than the wild-type enzyme. In the absence of the activator 3-PGA, the up-regulatory mutant ADPG-PPs described herein have less than 5% of the activity levels exhibited by the fully activated enzyme. In contrast, the *E. coli* glgC16 enzyme is essentially unregulated. Even in the total absence of the activator, it has 60% of the activity exhibited by the fully activated wild-type enzyme.

In addition to differences in allosteric regulation, there are differences in the structures of the bacterial and higher plant ADPG-PPs. The bacterial enzyme is encoded by a single gene, *glgC*, which codes for a subunit of 50 kD molecular weight that assembles to form a homotetramer. The higher plant enzyme is

composed of two large subunits and two small subunits. Each subunit type is encoded by a distinct gene. As described above, it is believed that the large and small subunits play different roles in enzyme catalysis and allosteric regulation. Under certain conditions the small subunit is able to form a homotetrameric enzyme but requires more than 24-fold greater levels of 3-PGA for activation than the normal heterotetrameric enzyme, which is composed of two large subunits and two small subunits. The large subunit is not capable of forming an active enzyme by itself. These results indicate that the small subunit plays more of a catalytic role, while the large subunit regulates the activity of the small subunit.

The present invention includes procedures for efficiently generating and identify mutations in the large subunit or small subunit of ADPG-PP that result in the formation of up-regulated ADPG-PPs. When such up-regulated mutant ADPG-PPs are expressed, for example, in leaves and/or the appropriate developing organs of transgenic cereal, potato, tomato, and other starch-accumulating plants, both starch production and yield are increased. By contrast, the Monsanto invention is only concerned with increase starch production.

The Monsanto approach involves the use of random mutagenesis techniques and screening for mutants by iodine staining to generate unregulated mutants. However, the Monsanto approach would not identify unregulated mutant enzymes because the approach provides no means to distinguish unregulated mutant enzymes from the wild-type enzyme based on iodine staining. Moreover, the Monsanto approach cannot generate unregulated enzymes by mutagenesis of plant ADPG-PP sequences, since the plant enzyme is composed of two subunit types, with the large subunit increasing the sensitivity of the small subunit to the allosteric effector molecules.

Definitions and Methods

The following definitions and methods are provided to better define the present invention and to guide those of ordinary skill in the art in the practice of the present invention. Unless otherwise noted, terms are to be understood according to conventional usage by those of ordinary skill in the relevant art. Definitions of common terms in molecular biology may also be found in Rieger *et al.*, *Glossary of Genetics: Classical and Molecular*, 5th edition, Springer-Verlag: New York, 1991; and Lewin, *Genes VI*, Oxford University Press: New York, 1997. The nomenclature for DNA bases as set forth at 37 C.F.R. § 1.822 is used. The standard one- and three-letter nomenclature for amino acid residues is used.

"Yield" refers to the amount of harvestable material, for instance, seeds or tubers. Yield is generally defined in units of mass.

"Productivity" refers to the total biomass produced by a plant, both harvestable and non-harvestable.

"Increased [characteristic] compared to a wild-type plant cell". When used in the context of plant characteristics (for instance, increased yield compared to the wild-type plant cell, increased productivity compared to the wild-type plant cell, increased starch production compared to the wild-type plant cell, increased size compared to the wild-type plant cell, increased rate of growth compared to the wild-type plant cell, or increased number of seeds compared to the wild-type plant cell) the word "increased" means that the mean value of a characteristic in a population, when compared with another (wild-type) population is measurably and statistically significantly greater than the mean value of the characteristic in question for the characteristic in the wild-type population.

"ADPG-PP nucleic acid". The term "ADPG-PP nucleic acid" refers to a native (or wild-type) nucleic acid

that encodes an ADPG-PP polypeptide (including, but not limited to, cDNA and genomic sequences) and fragments thereof.

"ADPG-PP Polypeptide". The term "ADPG-PP polypeptide" refers to a polypeptide encoded by an ADPG-PP nucleic acid. An ADPG-PP polypeptide can be produced by the expression of a recombinant ADPG-PP nucleic acid or can be chemically synthesized. Techniques for chemical synthesis of polypeptides are described, for example, in Merrifield, *J. Amer. Chem. Soc.* **85**:2149-2156, 1963.

"Mutant" or "mutated" ADPG-PP nucleic acid. "Mutated" or "mutant" ADPG-PP nucleic acids include a change in at least one base of the protein-coding region of a native ADPG-PP nucleic acid that results in a corresponding change in an amino acid of an ADPG-PP polypeptide encoded by the ADPG-PP nucleic acid. Included are insertions, deletions (including deletions of one or more nucleotides internal to the protein-coding region or deletions (or truncations) from one or both ends of the protein-coding region), and substitutions. Such mutated nucleic acids can be produced by any standard mutagenesis technique, including, but not limited to, those described in *Molecular Cloning: A Laboratory Manual*, 2nd ed., vol. 1-3, ed. Sambrook *et al.*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989 (hereinafter, "Sambrook, 1989"); *Current Protocols in Molecular Biology*, ed. Ausubel *et al.*, Greene Publishing and Wiley-Interscience, New York, 1994 (with periodic updates) (hereinafter, "Ausubel, 1994").

"Mutagenized". The word "mutagenized" is synonymous with the word "mutated" and means being altered from the native, natural state by any means including by chemical mutagens, electromagnetic radiation, genetic engineering, or molecular biology techniques. Thus a nucleic acid may be mutagenized by insertion,

substitution, or deletion of a nucleic acid, and a peptide may be mutagenized by insertion, substitution, or deletion of an amino acid residue such that the nucleic acid or peptide, respectively, is altered from its natural, native state.

"Native". The term "native" refers to a naturally-occurring ("wild-type") nucleic acid or polypeptide. The native nucleic acid or protein may have been physically derived from a particular organism in which it is naturally occurring or may be a synthetically constructed nucleic acid or protein that is identical to the naturally-occurring nucleic acid or protein.

"Isolated". An "isolated" nucleic acid is one that has been substantially separated or purified away from other nucleic acid sequences in the cell of the organism in which the nucleic acid naturally occurs, i.e., other chromosomal and extrachromosomal DNA and RNA, by conventional nucleic acid-purification methods. The term also embraces recombinant nucleic acids and chemically synthesized nucleic acids.

"Fragments, Probes, and Primers". A fragment of an ADPG-PP nucleic acid is less than full length and is capable of hybridizing specifically with a native ADPG-PP nucleic acid under stringent hybridization conditions. The length of such a fragment is at least 15, 20, 30, 40, or 50 nucleotides of a native ADPG-PP nucleic acid sequence or another target sequence, e.g., sequences flanking the cloning site of a vector in which an ADPG-PP sequence is cloned.

A "probe" is an isolated nucleic acid to which is attached a conventional detectable label or reporter molecule, *e.g.*, a radioactive isotope, ligand, chemiluminescent agent, or enzyme. "Primers" are isolated nucleic acids that are annealed to a complementary target DNA strand by nucleic acid hybridization to form a hybrid between the primer and the target DNA strand, then extended along the target DNA strand by a polymerase, *e.g.*, a DNA polymerase.

Primer pairs can be used for amplification of a nucleic acid sequence, *e.g.*, by the polymerase chain reaction (PCR) or other conventional nucleic-acid amplification methods.

Probes and primers are generally 15 nucleotides or more in length, and maybe 20, 25, or 30 nucleotides or more. Such probes and primers hybridize specifically to a target nucleic acid under high-stringency hybridization conditions.

Probes and primers have complete sequence similarity with the target sequence in at least about 7-15 consecutive nucleotides to permit hybridization under high stringency conditions, although probes differing from a target sequence and that retain the ability to hybridize to the target sequence may be designed by conventional methods and are useful for introducing nucleotide sequence mutations and corresponding amino acid sequence mutations.

Methods for preparing and using probes and primers are described, for example, in Sambrook, 1989; Ausubel, 1994; and Innis *et al.*, *PCR Protocols: A Guide to Methods and Applications*, Academic Press: San Diego, 1990. PCR-primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5©, 1991, Whitehead Institute for Biomedical Research, Cambridge, MA).

"Portions" of DNA. In relation to methods of the invention directed to deletion of DNA sequences that correspond to parts of the N- or C- terminus of the ADPG-PP protein, the term "portion" is used to indicate a number of contiguous nucleotides, from either the 3' or the 5' terminus of the DNA, corresponding to 5, 10, 15, 17, 25, 30, 40 or more amino acid residues from the N- or C- terminal. Deletion may be performed by standard molecular biology techniques (Sambrook et al., 1989)

"Substantial Similarity". A first nucleic acid is "substantially similar" to a second nucleic acid if, when optimally aligned (with appropriate nucleotide insertions or deletions) with the other nucleic acid (or its complementary strand), there is at least about 50%, 60%, 70%, 75%, 80%, 85%, 90% or 95% identity. Sequence similarity can be determined by comparing the nucleotide sequences of two nucleic acids using sequence analysis software such as the BLAST sequence analysis software available from the NCBI. A particularly useful tool is BLAST 2.0 program "gapped blastn" set to default parameters. Another useful software product is the Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, Madison, WI.

Alternatively, two nucleic acids are substantially similar if they hybridize under stringent conditions, as defined below.

"Operably Linked". A first nucleic-acid sequence is "operably" linked with a second nucleic-acid sequence when the first nucleic-acid sequence is placed in a functional relationship with the second nucleic-acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA sequences are contiguous and, where necessary to join two protein coding regions, in reading frame.

"Recombinant". A "recombinant" nucleic acid is made by an artificial combination of two otherwise separated segments of sequence, *e.g.*, by chemical synthesis or by the manipulation of isolated segments of nucleic acids by genetic engineering techniques.

Techniques for nucleic-acid manipulation are well-known (*see, e.g.*, Sambrook, 1989, and Ausubel, 1994). Methods for chemical synthesis of nucleic acids are discussed, for example, in Beaucage and Carruthers,

Tetra. Letts. **22**:1859-1862, 1981, and Matteucci *et al.*, *J. Am. Chem. Soc.* **103**:3185, 1981. Chemical synthesis of nucleic acids can be performed, for example, on commercial automated oligonucleotide synthesizers.

Preparation of Recombinant or Chemically Synthesized Nucleic acids; Vectors, Transformation, Host cells. Natural or synthetic nucleic acids according to the present invention can be incorporated into recombinant nucleic-acid constructs, typically DNA constructs, capable of introduction into and replication in a host cell. Such a construct may be a vector that includes a replication system and sequences that are capable of transcription and translation of a polypeptide-encoding sequence in a given host cell.

For the practice of the present invention, conventional compositions and methods for preparing and using vectors and host cells are employed, as discussed, *inter alia*, in Sambrook, 1989, or Ausubel, 1994.

A cell, tissue, organ, or organism into which a foreign nucleic acid has been introduced is considered "transformed", "transfected", or "transgenic." A "transgenic" or "transformed" cell or organism also includes progeny of the cell or organism, including progeny produced from a sexual cross that includes the ADPG-PP transgene from one or both parents.

Nucleic-acid constructs, or vectors, for use with prokaryotic or eukaryotic hosts, include a nucleic acid sequence that encodes an ADPG-PP polypeptide or a portion thereof and other vector sequences known in the art and appropriate for a given host cell, including, but not limited to, well known transcription and translation-initiation sequences; an origin of replication or autonomously replicating sequence (ARS); expression control sequences, including, but not limited to, promoter and enhancer sequences; processing information sites, such as ribosome-binding sites, RNA splice sites, polyadenylation sites, transcriptional

terminator sequences, and mRNA stabilizing sequences; secretion, transit, and other peptide sequences that allow the protein to cross and/or lodge in a cell membrane or be secreted from a cell; selectable or screenable marker genes, etc.

Nucleic-acid constructs can be introduced into a host cell by any suitable conventional method, including electroporation; transfection employing calcium chloride, rubidium chloride calcium phosphate, DEAE-dextran, or other substances; microprojectile bombardment; lipofection; infection (where the vector is an infectious agent; etc.) See, e.g., Sambrook, 1989, and Ausubel, 1994.

Nucleic acid constructs that express a mutated ADPG-PP according to the invention can be introduced into a variety of host cells or organisms in order to alter starch biosynthesis by the cell or organism, particularly higher plant cells, but also including other prokaryotic or eukaryotic host cells that synthesize starch.

Any well known vector suitable for stable transformation of plant cells and/or for the establishment of transgenic plants may be used, including those described in, *e.g.*, Pouwels *et al.*, *Cloning Vectors: A Laboratory Manual*, 1985, supp. 1987); Weissbach and Weissbach, *Methods for Plant Molecular Biology*, Academic Press, 1989; and Gelvin *et al.*, *Plant Molecular Biology Manual*, Kluwer Academic Publishers, 1990. Such plant expression vectors can include expression control sequences (*e.g.*, inducible or constitutive, environmentally or developmentally regulated, or cell- or tissue-specific expression-control sequences).

Examples of constitutive plant promoters useful for expressing ADPG-PP in plants include, but are not limited to, the cauliflower mosaic virus (CaMV) 35S promoter (*see, e.g.*, Odel *et al.*, *Nature* **313**:810, 1985;

Dekeyser *et al.*, *Plant Cell* **2**:591, 1990; and Terada and Shimamoto, *Mol. Gen. Genet.* **220**:389, 1990); the nopaline synthase promoter (An *et al.*, *Plant Physiol.* 88:547, 1988) and the octopine synthase promoter (Fromm *et al.*, *Plant Cell* 1:977, 1989).

A variety of plant gene promoters that are regulated in response to environmental, hormonal, chemical, and/or developmental signals, also can be used for protein expression in plant cells, including promoters regulated by (1) heat (Callis *et al.*, *Plant Physiol.* 88:965, 1988), (2) light (*e.g.*, pea rbcS-3A promoter, Kuhlemeier *et al.*, *Plant Cell* **1**:471, 1989; maize rbcS promoter, Schaffner and Sheen, *Plant Cell* **3**:997, 1991; or chlorophyll a/b-binding protein promoter, Simpson *et al.*, *EMBO J.* **4**:2723, 1985), (3) hormones, such as abscisic acid (Marcotte *et al.*, *Plant Cell* **1**:969, 1989), (4) wounding (*e.g.*, *wunI*, Siebertz *et al.*, *Plant Cell* **1**:961, 1989); or (5) chemicals such as methyl jasminate, salicylic acid, or a safener. It may also be advantageous to employ (6) organ-specific promoters (*e.g.*, Roshal *et al.*, *EMBO J.* **6**:1155, 1987; Schernthaner *et al.*, *EMBO J.* **7**:1249, 1988; Bustos *et al.*, *Plant Cell* **1**:839, 1989; Zheng *et al.*, *Plant J.* 4:357-366, 1993).

Plant expression vectors can include regulatory sequences from the 3'-untranslated region of plant genes (Thornburg *et al.*, *Proc. Natl. Acad. Sci. USA* **84**:744 (1987); An *et al.*, *Plant Cell* **1**:115 (1989), *e.g.*, a 3' terminator region to increase mRNA stability of the mRNA, such as the PI-II terminator region of potato or the octopine or nopaline synthase 3' terminator regions.

Useful dominant selectable marker genes for expression in plant cells include, but are not limited to: genes encoding antibiotic resistance genes (*e.g.*,

resistance to hygromycin, kanamycin, bleomycin, G418, streptomycin or spectinomycin); and herbicide resistance genes (*e.g.*, phosphinothricin acetyltransferase). Useful screenable markers include, but are not limited to, β-glucuronidase and green fluorescent protein.

Nucleic-Acid Hybridization. The nucleic-acid probes and primers of the present invention hybridize under stringent conditions to a target DNA sequence. Nucleic acid hybridization is discussed in Sambrook, 1989, at 9.52-9.55. *See also*, Sambrook, 1989 at 9.47-9.52, 9.56-9.58; Kanehisa, *Nucl. Acids Res.* 12:203-213, 1984; and Wetmur and Davidson, *J. Mol. Biol.* **31**:349-370, 1968.

Nucleic acid and protein sequence, similarity, homology, and identity. The similarity between two nucleic acid sequences, or two amino acid sequences, is expressed in terms of sequence identity. Sequence identity is frequently measured in terms of percentage identity (or similarity or homology); the higher the percentage, the more similar the two sequences.

Methods of alignment of sequences for comparison are well known in the art. Various programs and alignment algorithms may be found, for instance, in: Smith and Waterman (1981); Needleman and Wunsch (1970); Pearson and Lipman (1988); Higgins and Sharp (1988); Higgins and Sharp (1989); Corpet et al. (1988); Huang et al. (1992); and Pearson et al. (1994). Altschul et al. (1994) presents a detailed consideration of sequence alignment methods and homology calculations.

A particularly useful tool is the NCBI Basic Local Alignment Search Tool (BLAST) (Altschul et al., *J. Mol. Biol.* **215**:403-410, 1990) which is available from the National Center for Biotechnology Information (NCBI, Bethesda, MD) and on the Internet. It can be accessed at http://www.ncbi.nlm.nih.gov/BLAST/. A description of how to determine sequence identity using this program is

available at
http://www.ncbi.nlm.nih.gov/BLAST/blast_help.html.

Nucleic acid sequence similarity can be determined by using the NCBI BLAST 2.0 "gapped blastn" program set to default parameters. Amino acid sequence similarity can be compared using NCBI 2.0 "gapped blastP" set to default parameters. The "gapped" feature of these programs allows gaps (deletions and insertions) to be introduced into the sequences to be aligned. Allowing such gaps means that similar regions are not broken into several separate segments, allowing a truer determination of sequence identity (and therefore similarity) to be made.

Another useful BLAST tool is the Position-Specific Iterated BLAST (PSI-BLAST) that provides a sensitive way to look for sequence homologues. The program first performs a gapped BLAST database search. The PSI-BLAST program uses the information from any significant alignments returned to construct a position-specific score matrix, which replaces the query sequence for the next round of database searching. PSI-BLAST may be iterated until no new significant alignments are found.

Protein Orthologs. Orthologs of a particular protein are typically characterized by possession of at least 50% sequence identity counted over the full length alignment with the amino acid sequence of the gene in question using the NCBI Blast 2.0, gapped blastp set to default parameters. Proteins with even greater similarity to the reference sequences will show increasing percentage identities when assessed by this method, such as at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 90%, or at least 95% sequence identity. When less than the entire sequence is being compared for sequence identity, homologs will typically possess at least 75% sequence identity over short windows of 10-20 amino acids, and may possess sequence identities of at least 85% or at least 90% or 95%, depending on their similarity to the

reference sequence. Methods for determining sequence identity over such short windows are described at http://www.ncbi.nlm.nih.gov/BLAST/blast_FAQs.html.

One of ordinary skill in the art will appreciate that these sequence identity ranges are provided for guidance only; it is entirely possible that strongly significant homologs could be obtained that fall outside of the ranges provided. The present invention provides not only the peptide homologs are described above, but also nucleic acid molecules that encode such homologs.

Hybridization and Stringent Conditions. An alternative indication that two nucleic acid molecules are closely related is that the two molecules hybridize to each other under stringent conditions. Stringent conditions are sequence dependent and are different under different environmental parameters. Generally, stringent conditions are selected to be about 5°C to 20°C lower than the thermal melting point (Tm) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Conditions for nucleic acid hybridization and calculation of stringencies can be found in Sambrook et al. (1989) and Tijssen (1993). Nucleic acid molecules that hybridize under stringent conditions to the target sequences will typically hybridize to a probe based on either the entire target cDNA or selected portions of the cDNA under wash conditions of 0.2x SSC, 0.1% SDS at 65°C. Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences, due to the degeneracy of the genetic code. It is understood that changes in nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequence that all encode substantially the same protein.

Regarding the amplification of a target nucleic-acid sequence (*e.g.*, by PCR) using a particular amplification primer pair, "stringent conditions" are

conditions that permit the primer pair to hybridize only to the target nucleic-acid sequence to which a primer having the corresponding wild-type sequence (or its complement) would bind and preferably to produce a unique amplification product.

The term "specific for (a target sequence)" indicates that a probe or primer hybridizes under given hybridization conditions only to the target sequence in a sample comprising the target sequence.

Nucleic-Acid Amplification. As used herein, "amplified DNA" refers to the product of nucleic-acid amplification of a target nucleic-acid sequence. Nucleic-acid amplification can be accomplished by any of the various nucleic-acid amplification methods known in the art, including the polymerase chain reaction (PCR). A variety of amplification methods are known in the art and are described, *inter alia*, in U.S. Patent Nos. 4,683,195 and 4,683,202, and in *PCR Protocols: A Guide to Methods and Applications*, ed. Innis *et al.*, Academic Press, San Diego, 1990.

Mutagenesis of ADPG-PP Nucleic Acids. Using ADPG-PP nucleic acids that are known in the art or that are isolated using such well-known nucleic acids, any conventional mutagenesis method can be used to mutagenize (i.e., artificially alter from its wild-type form) ADPG-PP nucleic acids, including chemical mutagenesis, oligonucleotide site-directed mutagenesis, chemical synthesis of a mutant ADPG-PP sequence, etc., resulting in substitutions, insertions, deletions, or combinations thereof. Nucleic acids so produced are called "mutant" nucleic acids.

Mutant ADPG-PP Polypeptides. A mutant ADPG-PP polypeptide is a polypeptide produced by the expression of a mutant nucleic acid, or a peptide that has been intentionally artificially altered from its wild-type form.

"Silent" mutations include substitutions of one or more base pairs that result in no change in the amino acid sequence of the polypeptide encoded by the sequence. "Conservative" mutations result in a conservative amino acid substitution in one or more amino acid residues of the polypeptide encoded by the nucleic-acid sequence. Examples of conservative amino acid substitutions include: Ser for Ala; Lys for Arg; Gln or His for Asn; Glu for Asp; Ser for Cys; Asn for Gln; Asp for Glu; Pro for Gly; Asn or Gln for His; Leu or Val for Ile; Ile or Val for Leu; Arg, Gln, or Glu for Lys; Leu or Ile for Met; Met, Leu, or Tyr for Phe; Thr for Ser; Ser for Tyr; Tyr for Trp; Trp or Phe for Tyr; and Ile or Leu for Val.

Polypeptide Sequence Homology. ADPG-PP polypeptides encompassed by the present invention are at least about 70%, 80%, 90%, or 95% homologous to a native ADGP-PP polypeptide.

Polypeptide homology can be analyzed by conventional methods, e.g., using sequence analysis software such as the Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, Madison, WI, or the NCBI BLAST 2.0 blast software. Polypeptide sequence analysis software matches homologous sequences using measures of homology assigned to various substitutions, deletions, substitutions, and other modifications.

"Isolated," "Purified," "Homogeneous" Polypeptides. An "isolated" polypeptide is separated from the cellular components (nucleic acids, lipids, carbohydrates, and other polypeptides) that naturally accompany it. Such a polypeptide can also be referred to as "pure" or "homogeneous" or "substantially" pure or homogeneous. Thus, a polypeptide which is chemically synthesized or recombinant (i.e., the product of the expression of a recombinant nucleic acid, even if expressed in a homologous cell type) is considered to be isolated. A monomeric polypeptide is isolated when at least 60% by

weight of a sample is composed of the polypeptide. In some cases it is advantageous to have 90%, 95%, or 99% polypeptide by weight in a sample. Protein purity or homogeneity is indicated, for example, by polyacrylamide gel electrophoresis of a protein sample, followed by visualization of a single polypeptide band upon staining the polyacrylamide gel; high performance liquid chromatography; or other conventional methods.

Protein Purification. Polypeptides according to the present invention can be purified by any conventional method. *See, e.g., Guide to Protein Purification*, ed. Deutscher, *Meth. Enzymol.* **185**, Academic Press, San Diego, 1990; and Scopes, *Protein Purification: Principles and Practice*, Springer Verlag, New York, 1982.

ADPG-PP "Biological Activity" or "Enzyme Activity". The terms "biological activity", "enzyme activity", "biologically active", "activity", and "active" refer primarily to the characteristic biological activity or activities of a native ADPG-PP polypeptide, including, but not limited to, catalyzing the initial step in α-glucan and starch synthesis: glucose-1-phosphate + ATP --> ADP-glucose + PPi. Other activities include allosteric regulation by 3-PGA and Pi.

"Allosteric". The word allosteric literally means "another site". An enzyme is said to be "allosteric" or "allosterically regulated" if its enzymatic activity is regulated by the binding of non-substrate molecules at a site other than the active site of the enzyme, i.e., an "allosteric site". The binding of non-substrate molecules at allosteric sites effects the binding kinetics of the substrate-binding (active) site. A molecule that increases binding or decreases dissociation of an enzyme and a substrate complex, thus increasing enzyme activity, is called an "activator" molecule. A molecule that decreases binding or increases dissociation of an enzyme and a substrate complex, thus

decreasing enzyme activity, is called an "repressor" or "inhibitor" molecule.

"Up-regulated" (or "up-regulatory"). A mutant enzyme is "up-regulated" or "up-regulatory" when that enzyme has a higher biological activity than the wild-type enzyme under physiological conditions. The up-regulated enzyme may have at least about 10, 20, 40, 60, 80, or 100% higher biological activity than wild-type. Under some circumstances, the up-regulated mutant may have greater than 100% higher biological activity than the wild-type. An up-regulated mutant exhibits increased biological activity by virtue of its mutation(s). The effect of the mutation(s) may be to inherently increase biological activity independent of allosteric regulation (for instance by increasing binding or decreasing dissociation of the enzyme-substrate complex), to increase sensitivity to an allosteric activator (e.g., 3-PGA), and/or to decrease sensitivity to an allosteric inhibitor (e.g., Pi). Any such effect brought about by a mutation will result in an increase in enzyme activity making the mutant enzyme up-regulated.

Fusion Polypeptides. The present invention also provides fusion polypeptides including, for example, heterologous fusion polypeptides in which an ADPG-PP sequence is joined to a well-known fusion partner. Such fusion polypeptides can exhibit biological properties (such as substrate or ligand binding, enzymatic activity, antigenic determinants, *etc.*) derived from each of the fused sequences. Fusion polypeptides are preferably made by standard recombinant DNA techniques.

Transgenic Plants; Plant Transformation and Regeneration

Various nucleic acid constructs that include a mutant ADPG-PP nucleic acid are useful for producing transgenic plants with altered starch biosynthesis, increased starch production, and increased yield. Such plants may also grow faster and produce more seeds.

ADPG-PP nucleic acids can be expressed in plants or plant cells under the control of a suitable operably linked promoter that is capable of expression in a cell of a particular plant. Any well-known method can be employed for plant cell transformation, culture, and regeneration in the practice of the present invention with regard to a particular plant species. Conventional methods for introduction of foreign DNA into plant cells include, but are not limited to: (1) *Agrobacterium*-mediated transformation (Lichtenstein and Fuller In: *Genetic Engineering*, Vol 6, Rigby, ed., London, Academic Press, 1987; and Lichtenstein and Draper, in: *DNA Cloning*, Vol II, Glover, ed., Oxford, IRI Press, 1985); (2) particle delivery (*see, e.g.*, Gordon-Kamm *et al.*, *Plant Cell* **2**:603, 1990; or BioRad Technical Bulletin 1687); (3) microinjection (*see, e.g.*, Green *et al.*, *Plant Tissue and Cell Culture*, Academic Press, New York, 1987); (4) polyethylene glycol (PEG) procedures (*see, e.g.*, Draper *et al.*, *Plant Cell Physiol.* **23**:451, 1982); Zhang and Wu, *Theor. Appl. Genet.* **76**:835, 1988); (5) liposome-mediated DNA uptake (*see, e.g.*, Freeman *et al.*, *Plant Cell Physiol.* **25**:1353, 1984); (6) electroporation (*see, e.g.*, Fromm *et al.*, *Nature* **319**:791 (1986)); and (7) vortexing methods (*see, e.g.*, Kindle, *Proc. Natl. Acad. Sci. USA* **87**:1228 (1990)).

The term "plant" encompasses any higher plant and progeny thereof, including monocots (*e.g.*, corn, rice, wheat, barley, *etc.*), dicots (*e.g.*, potato, tomato, *etc.*), and includes parts of plants, including seeds, fruit, tubers, *etc.*

The invention will be better understood by reference to the following Examples. The scope of the invention is not to be considered limited thereto.

EXAMPLES

EXAMPLE 1: Generation and Identification of ADPG-PP Mutants

Bacterial mutants that lack functional ADPG-PP activity, e.g., *glgC⁻ E. coli* strains such as AC70R1-504 (Leung *et al.*, *J. Bacteriol.* **167**:82-87, 1986) are useful for studying the function of mutant plant ADPG-PP. Co-expression of cDNAs encoding the large and small subunits of plant ADGP-PP in such bacterial mutants results in production of a plant ADPG-PP enzyme that is capable of complementing the *glgC⁻* mutation, restoring glycogen production. Glycogen accumulation can be easily scored by exposing the cells to I_2 vapor. Within 30 sec, cells co-expressing wild-type plant ADGP-PP stain purplish-black, whereas the untransformed *glgC⁻* host strain stains light yellow.

The ability of the plant enzyme to complement complement *glgC⁻ E. coli* mutants was exploited in several related mutagenesis approaches that were developed to generate and identify up-regulated and other ADPG-PP mutants. One of the plant ADPG-PP subunit cDNA sequences was mutagenized then co-expressed with the unmutated counterpart subunit.

For example, in one approach, the large subunit is mutagenized, e.g., using hydroxylamine, and cells co-expressing a mutagenized large subunit and a wild-type small subunit were screened for mutants defective in glycogen accumulation by I_2 staining (Greene *et al.*, *Proc. Natl. Acad. Sci. USA* **93**:1509-1513, 1996; Greene *et al.*, *Plant Physiol.* **112**:1315-1320, 1996). Mutant ADPG-PP enzymes defective for allosteric regulation were then identified by assaying crude extracts of non-staining cells under saturating conditions of 3-PGA (10 mM, or 100-fold greater levels than required to activate the enzyme by 50%).

Hydroxylamine causes G/C-to-A/T transitions. Mutations can also be introduced, for example, by various other conventional mutagenesis techniques, e.g., by using other chemical mutagens or by PCR performed under limiting substrate and co-factor conditions. Site-directed mutagenesis may also be used to introduce insertions, deletions or substitutions. Various standard techniques are known to carry out site-directed mutagenesis (Sambrook et al., 1989), and commercial kits are also available such as the QUICKCHANGE™ mutagenesis site-directed mutagenesis kit (STRATEGENE™, CA).

Plasmid DNA encoding the large subunit was mutagenized at 37°C for 24 h in the presence of 0.8 M hydroxylamine-HCl, 50 mM sodium phosphate, pH 6.0, and 1 mM EDTA (Isackson and Bertrand, *Proc. Natl. Acad. Sci. USA* **82**:6226-6230, 1985). The reaction was terminated by the addition of 50 μL of 1 M Tris base, ethanol precipitated, and resuspended in 10 mM Tris-Cl, pH 8.0, and 1 mM EDTA, pH 8.0. DNA was then transformed into *E. coli* AC70R1-504 (Leung *et al.*, *J. Bacteriol.* **167**:82-87, 1986), a *glgC*⁻ strain, carrying pML7, the wild-type small subunit cDNA (the construction of pML7 is described below). After overnight growth on Luria broth (LB) media, colonies were then screened for the production of glycogen by replica plating the colonies onto Kornberg media containing 2% glucose (Greene *et al.*, *Proc. Natl. Acad. Sci. USA* **93**:1509-1513, 1996; Greene *et al.*, *Plant Physiol.* **112**:1315-1320, 1996), growing the cells overnight, then exposing the plates to I_2 vapor. Glycogen-deficient cells stain a faint yellow, whereas glycogen-positive cells stain a purplish-black color. Cells that were deficient for glycogen accumulation were grown overnight in 1 mL of LB media. Cells were collected by brief centrifugation, lysed in 100 μL of sucrose lysis buffer by repeated freeze/thawing, then clarified by centrifugation. The cell extracts were

analyzed by ELISA using rabbit monospecific anti-large subunit antibodies and goat anti-rabbit antibodies linked to horseradish peroxidase (Greene *et al.*, *Proc. Natl. Acad. Sci. USA* **93**:1509-1513, 1996). Cells that accumulated normal antigen levels for both the large and small subunit were then analyzed by enzyme assay. Cells were grown overnight on an enriched solid media at 37°C then lysed by sonication in 250 μL of a sucrose buffer (50 mM Hepes, pH 8.0, 20% sucrose, 5 mM magnesium chloride ($MgCl_2$), 1 mM EDTA, 5 mM dithiothreitol (DTT), 5 mM sodium/potassium phosphate buffer, pH 8.0, containing 500 μg/mL lysozyme, 0.5 μg/mL pepstatin, 0.5 μg/mL leupeptin, 0.5 mM benzamidine, and 1 mM phenyl methyl sulfonyl fluoride. After centrifugation, crude extracts were assayed for ADPG-PP activity in the pyrophosphorylysis direction as described (Greene *et al.*, *Proc. Natl. Acad. Sci. USA* **93**:1509-1513, 1996; Greene *et al.*, *Plant Physiol.* **112**:1315-1320, 1996), except that the reaction contained 20 mM 3-PGA. Under these conditions, several mutant lines could be identified that contained normal levels of ADPG-PP enzyme activity when assayed under saturating activator conditions (Greene *et al.*, *Proc. Natl. Acad. Sci. USA* **93**:1509-1513, 1996; Greene *et al.*, *Plant Physiol.* **112**:1315-1320, 1996). This was verified for mutant 345, which displayed an $A_{0.5}$ (defined as the amount of 3-PGA required for 50% activation) of 4.0 mM or about 30- to 40-fold greater than the wild-type enzyme (Greene *et al.*, *Proc. Natl. Acad. Sci. USA* **93**:1509-1513, 1996).

To substantiate that the mutation affects only the allosteric regulatory properties of the enzyme, ADPG-PP was purified by biochemical chromatography techniques, and the activity of the mutant enzyme was investigated. Cells were grown in 850 mL of LB in Fernbach flasks and induced by the addition of IPTG and nalidixic acid. Large scale crude extracts were obtained by lysing and

clarifying as described above. The crude extract was then subjected to a 33-55% ammonium sulfate fractionation. The resulting precipitate was resuspended in sucrose buffer and desalted by dialysis. The desalted sample (30 mg) was diluted to 30 mL in buffer A (5 mM potassium phosphate, pH 7.5, 50 mM glycylglycine, pH 7.5, 5 mM $MgCl_2$, 1 mM EDTA, and 5 mM DTT) and was applied to a Millipore MEMSEP 1010 DEAE-chromatography column previously equilibrated with buffer A. Enzyme was eluted from the column by a 200 mL linear gradient of buffer B (50 mM potassium phosphate, pH 6.0, 5 mM $MgCl_2$, 1 mM EDTA, 5 mM DTT, and 400 mM potassium chloride). Fractions containing enzyme activity were pooled, concentrated by addition of ammonium sulfate to 75%, resuspended in sucrose buffer, and desalted by dialysis. Aliquots of enzyme concentrate were stored at -80°C until needed for enzyme studies. Kinetic parameters, i.e., K_m values for glucose 1-phosphate, ATP, and Mg^{2+} and $A_{0.5}$ for 3-PGA and $I_{0.5}$ for Pi, were defined with the synthesis (forward) assay, which measures the incorporation of ^{14}C-glucose-1-phosphate into ADPglucose.

Once a mutant enzyme was identified as defective in allosteric regulation, the mutated plasmid DNA coding for the large subunit coding for the large subunit was isolated and subjected to a second round of hydroxylamine mutagenesis and re-transformed into AC70R1-504 harboring a plasmid DNA containing the small subunit counterpart cDNA. The mutated plasmid DNA was isolated then transformed into *E. coli*. Cells containing the plasmid DNA of interest were then selected by growth on enriched media containing the appropriate antibiotic. After replica plating and growth overnight, revertant cells that accumulated glycogen were identified by I_2 staining. The enzyme activities from these revertant cells were partially purified and characterized kinetically as described above. The nature of the second-site mutation (because

hydroxylamine causes G/C to A/T transitions, the reversion event is caused by second-site mutation) was identified by DNA sequencing of the cDNA sequence.

The effect of this second-site mutation on enzyme function was then evaluated by introducing the second-site mutation by itself (i.e., without the original mutation) into a wild-type cDNA sequence for the appropriate subunit by site-directed mutagenesis (using a commercial site-directed mutagenesis kit (such as the Quickchange™ kit from Stratagene, CA) according to the manufacturer's instructions). The resulting subunit cDNA including the second-site mutation was then transformed into and co-expressed in AC70R1-504 together with the counterpart wild-type subunit. The resulting mutant ADPG-PP enzyme was then characterized.

Because of its specificity towards deoxycytidine nucleosides, the spectrum of mutations generated by hydroxylamine is limited. Mutations that result in the replacement of Lys and Tyr residues, which are known to play a prominent role in the bacterial enzyme, are not expected. To generate as wide a spectrum of mutations as possible, a PCR strategy or other conventional site-directed mutagenesis method can be employed. For example, the large subunit or small subunit sequences can be amplified under conditions (high Mn^{2+} or limiting nucleotide concentrations) that result in the misincorporation of nucleotides (Erlich, In: *PCR Technology -- Principles and Applications for DNA Amplification*, New York: Stockton Press, 1989; Slepak *et al.*, *J. Biol. Chem.* **268**:1414-1423, 1993). The amplified sequences can then be purified, cleaved with the appropriate restriction enzymes, cloned into the appropriate sites of the expression vectors, and expressed in *E. coli*. Mutants defective in pyrophosphorylase activity can then be screened initially by I_2 staining, followed by ELISA and enzyme assays as described above.

In addition to the protocol described above, up-regulated enzymes can be identified directly by using a variant of the mutagenesis and selection protocol described above. Cells expressing wild-type ADPG-PP enzyme on medium or high copy number plasmids are close to saturating the amount of glycogen that can be accumulated under normal culture conditions. Therefore, cells expressing an up-regulated ADPG-PP cannot be easily distinguished from cells expressing the wild-type enzyme by iodine staining.

To distinguish the expression of an up-regulated enzyme from the expression of the wild-type enzyme, the large or small subunit cDNA that will be mutagenized is cloned into a low copy number plasmid such as pWSK 28 or other plasmids described in Wang and Kushner, *Gene* **100**:195-199, 1991 (Other low copy-number plasmids may be obtained commercially from such companies as Invitrogen, Pharmacia and New England Biolabs. Low-copy-number vectors are generally larger than high-copy-number vectors, having sizes of 6 or 7 kilobases or more. These low-copy-number plasmids contain, for example, a pSC101 origin of replication, which results in substantially fewer copies (*e.g.*, about 6 copies/cell) than a vector possessing a pBR325, Col E1, or ACYC origin of replication. As a result of the decrease in copy number, cells expressing wild-type large subunit sequences that are carried on a low-copy-number plasmid stain very lightly with I_2 when co-expressed with the small subunit sequences, providing a cleaner background to identify up-regulated mutants. Screening several thousand cells results in the isolation of several putative up-regulated mutants.

In addition to up-regulated mutations, other mutations can produce similar phenotypes, including mutations that increase plasmid copy number, promoter activity, and ribosome-binding efficiency. These other mutations can be identified by evaluating cellular

extracts by immunological techniques as described above, since it is expected that cells containing up-regulated mutants of ADPG-PPs contain the same amount of antigen as cells expressing the wild-type enzyme.

Alternatively, other types of mutations can be eliminated by subjecting the cDNA sequences alone to the chemical mutagen or PCR mutagenesis protocol (rather than the vector as a whole), then inserting the mutated cDNA sequences into an appropriate expression vector.

An alternative approach to identify up-regulatory mutants directly is by altering the level of glucose in the Kornberg media (normally 2%) to a level at which glycogen accumulation is approximately directly proportional to ADPG-PP activity (e.g., 0-0.5%). Cells expressing an up-regulated ADPG-PP mutant enzyme accumulate more glycogen at 0.1% glucose than cells expressing wild-type enzyme (*see* FIG. 3).

A third strategy for the generation of up-regulatory mutants is to mutate specific DNA sequences that code for specific peptide regions of either the large or small subunit. For example, peptide sequences at both the N- and C-terminus are required for normal allosteric regulation. Mutations in this region alter the allosteric response of the resulting enzyme (Ball and Preiss, *J. Biol. Chem.* **269**:24706-24711, 1994; Greene *et al.*, *Proc. Natl. Acad. Sci. USA* **93**:1509-1513, 1996; Greene *et al.*, *Plant Physiol.* **112**:1315-1320, 1996; Morell *et al.*, *J. Biol. Chem.* **263**:633-637, 1988). Other peptide regions of the large and small subunits may also be mutagenized to identify residues that play a role in allosteric regulation. Mutations can be generated randomly using PCR techniques under conditions that cause misincorporation of nucleotides or by using synthetic oligonucleotides that include random mutations. Mutations can also be generated by deletion of specific DNA sequences, especially those encoding

amino acid residues located at the N- or C-terminus, for example.

EXAMPLE 2: Generation and Identification of Up-Regulatory Mutants of the Higher Plant ADPG-PP

The expression plasmids, pMON17335 and pMON17336 (Inglesias *et al.*, *J. Biol. Chem.* **268**:1081-1086, 1993), that contain the small and large subunit cDNAs, respectively, were modified by incorporating double translation termination codons using a PCR approach. The primers shown in SEQ ID Nos. 13 and 14 were used to amplify a 538 bp DNA fragment covering the 3' end of the small subunit coding sequence. After digestion with KpnI and SacI, the amplified product was cloned into pMON17335 to produce pML5. The primers shown in SEQ ID NOS. 15 and 16 were used to amplify a 608 bp DNA fragment covering the 3' end of the large subunit coding sequence. After digestion with NheI and HindIII, the amplified product was cloned into pMON17336 to produce pML7 (FIG. 1). pML5-encoded small subunit polypeptide lacked 10 amino acid residues of the N-terminus of the mature protein. These ten amino acids were restored, together with an additional Met-Ala, by PCR using the primer shown in SEQ ID No. 17, which introduces a unique NcoI site (underlined DNA sequences) at the Met codon. The oligonucleotide primer shown in SEQ ID. No. 18, which spans a unique Kpn I site (DNA sequences underlined) located about 850 nucleotides from the newly created AUG codon, was employed together with the upstream oligonucleotide (SEQ ID No. 17) for PCR amplification. The amplified DNA fragment was digested with NcoI and KpnI and cloned into pML5 to give pML10 (shown in FIG. 2).

Random Mutagenesis. The large subunit plasmid, pMON17336, was subjected to hydroxylamine treatment and then transformed into the *glgC*$^-$ strain AC70R1-504 containing pMON17335 as described above. After

overnight growth, cells were replica-plated onto Kornberg media containing 2% glucose, grown for at least 12 hours, then exposed to iodine vapor. The glycogen-deficient cells were then analyzed by ELISA (Greene *et al., Proc. Natl. Acad. Sci. USA* **93**:1509-1513, 1996) and enzyme assays (*id.*) under saturating substrate and activator conditions. One glycogen-deficient mutant, mutant 345, which contained normal levels of enzyme activity, was further studied kinetically (*id.*). Fig. 5 shows substrate binding (K_ms) and 3-PGA activator ($A_{0.5}$) affinity properties of various ADPG-PPs: native (Sowokinos and Preiss, *Plant Physiol.* **69**:1459-1466, 1992); wild-type (Ballicora *et al., Plant Physiol.* **109**:245-251, 1995); mutant 345 (Greene *et al., Proc. Natl. Acad. Sci. USA* **93**:1509-1513, 1996); R20, UpReg-1, and ΔN17-LS (R20, UpReg-1, and ΔN17-LS are described below). The partially purified enzyme from mutant 345 displayed normal binding constants (K_m) for glucose-1-phosphate, ATP and Mg^{2+} (Fig. 5). However, the enzyme from mutant 345 required 28-fold greater levels of the activator 3-PGA than the wild-type enzyme. DNA sequence analysis indicated the presence of single base substitution that resulted in the replacement of a proline located at residue 52 by a leucine (Greene *et al., Proc. Natl. Acad. Sci. USA* **93**:1509-1513, 1996).

The large-subunit plasmid from mutant 345 was isolated and subjected to a second hydroxylamine treatment, co-expressed with pML10, and screened for glycogen production. Eight genetic revertants were isolated that displayed varying levels of glycogen accumulation as determined by I_2 staining. Four of the eight revertants, including R20, demonstrated wild-type staining phenotypes (Fig. 3) and higher than wild-type ADPG-PP activities (Class I). The remaining revertants showed either intermediate (Class II) or light staining (Class III) phenotypes.

Molecular characterization of the large subunit plasmid DNAs of Class I revertants R4, R10, R20, and R32 identified a single second-site mutation in close proximity to the primary mutation Pro52Leu in each of the revertants. R4 contained a base-pair mutation that replaced Pro at position 66 with a Leu, resulting in the addition of a second structural modification. R10 and R32 contained an identical second site mutation in which Gly at position 101 was replaced with Asn. In R20, a negatively charged Glu residue fourteen amino acids upstream from P52L was replaced with a positively charged Lys residue (SEQ ID Nos. 7 and 8); note that Lys38 is at position 40 in the recombinant cDNA sequence due to the addition of Met and Ala residues. Introduction of the positively charged Lys residue potentially enhances the enzyme's ability to interact with the negatively charged activator 3-PGA.

The R20 enzyme was partially purified to a specific activity of about 20 μmol/min/mg by differential ammonium sulfate precipitation, heat treatment, amino-propyl (C_3) chromatography, and DEAE anion-exchange chromatography and was estimated to be about 40% pure. SDS-polyacrylamide gel electrophoresis (SDS-PAGE) of the partially purified protein showed a dramatic increase in a broad band migrating in the 50-52 kD range. Immunoblot analysis showed that this band cross-reacted with monospecific IgG directed against the recombinant potato ADPG-PP large subunit or small subunit. Kinetic analysis of R20 enzyme showed that the values of K_m for glucose-1-phosphate, ATP, and Mg^{2+} were identical to the K_m of the wild-type recombinant enzyme (Fig. 5). The affinity of R20 enzyme for 3-PGA was increased about 38-fold compared to the P52L mutant, which displayed an $A_{0.5}$ of 117 μM, which is very close to that of the wild-type recombinant enzyme (Fig. 5).

Fig. 6 shows the results of studies to determine the Pi inhibition constants ($I_{0.5}$) of various ADPG-PPs: wild-type (Ballicora *et al.*, *Plant Physiol.* **109**:245-251,

1995); mutant 345 (Greene *et al.*, *Proc. Natl. Acad. Sci. USA* **93**:1509-1513, 1996); R20, UpReg-1, and AN17-LS. The $I_{0.5}$ value determined for the R20 ADPG-PP at 0.25 mM 3-PGA was roughly 4.5- and 8-fold higher than that of the native potato and wild-type recombinant enzymes, respectively (Fig. 6). The increase in positive charge in this region of the enzyme appears to give R20 ADPG-PP an increased affinity for 3-PGA, while decreasing its affinity for Pi.

Site-directed mutagenesis was conducted to introduce a Lys residue at position 38 (Glu38Lys) in pML7 that is independent of the P52L mutation. Of four putative Glu38Lys site-directed mutants (SEQ ID Nos. 1 and 2), all contained the specific mutation as determined by sequence analysis. Glu38Lys large-subunit plasmid DNA was then transformed into AC70R1-504 containing pML10 for purification and kinetic analysis. The Glu38Lys site-directed mutant was partially purified to a specific activity of 19.7 μmol/min/mg. Kinetic analysis identified an enzyme which was unchanged in its affinity for glucose-1-phosphate, ATP, and Mg^{2+} as compared to R20 (Fig. 5). Activation curves showed that the Glu38Lys mutant has an $A_{0.5}$ of 2 μM, a 58.5- and 80-fold increase in affinity for 3-PGA compared to R20 and wild-type recombinant ADPG-PP, respectively (Fig. 5) and a dramatic 2,250-fold increase in affinity for 3-PGA compared to the Pro52Leu mutant. The Glu38Lys site-directed mutant was also significantly less sensitive to Pi. Inhibition curves showed $I_{0.5}$ values of 0.6, 3.1, and 4.7 mM in the presence of 25, 125, and 250 μM 3-PGA. When compared to the $I_{0.5}$ values of the native potato and wild-type recombinant enzymes, the Glu38Lys mutant displayed a dramatic reduction in affinity for Pi (Fig. 6). The Glu38Lys mutation alone in the large subunit, when co-expressed with the unmutated small subunit, produced an up-regulated form of ADPG-PP by dramatically increasing the affinity for the activator and dramatically decreasing the affinity for the

inhibitor Pi. ADPG-PP containing the Glu38Lys mutant large subunit is an up-regulatory mutant. Henceforth, it is referred to as "UpReg-1" (SEQ ID Nos. 1 and 2). Phenotypically, the UpReg-1 mutant, when co-expressed with pML10 stains, is significantly darker on a glucose media compared to the wild-type recombinant and R20 (FIG. 3).

Additional site-directed mutagenesis of Glu 38 was conducted. An Ala residue is highly conserved at the analogous position on the small subunit. Substitution of Ala for Glu at position 38 generates an enzyme that is insensitive to activation by 3-PGA, similar to the Pro52Leu mutation (Greene *et al.*, *Proc. Natl. Acad. Sci. USA* **93**:1509-1513, 1996). An Arg residue resides at an analogous position in the *E. coli* ADPG-PP enzyme. Introduction of an Arg residue at position 38 generates an enzyme that remains as insensitive to 3-PGA activation as the Pro52Leu mutant. At high concentrations of 3-PGA (50 mM), an enzyme activity comparable to that of the wild-type recombinant enzyme can be achieved with the E38A and E38R mutants. The lack of activation in the E38R mutant at the lower concentrations of 3-PGA indicates that activation may not be strictly charge-dependent, but may also be affected by size constraints.

Similar studies were also conducted for R32 which contained a Gly101 Asn mutation (SEQ ID Nos. 9 and 10). Substrate and cofactor binding constants were also relatively unaffected for R32 AGPase (Fig. 5). R32 was more sensitive to 3-PGA activation, showing an $A_{0.5}$ of 414 μM, an 11-fold increase over the Pro52Leu mutant. Pi sensitivity was also lower in R32 AGPase (Figs. 6 and 7) but the $I_{0.5}$ values for this second-site revertant were more similar to the values determined for the WT recombinant enzyme than for the other second-site revertants. Thus, substitution of the polar amino acid Asn for the nonpolar Gly restored 3-PGA sensitivity in

R32 while only slightly changing the enzyme's affinity for Pi.

Site-directed studies were conducted to introduce an Asn at position 101 in the large subunit which was then co-expressed with the wildtype small subunit in bacterial cells. The resulting enzyme, UpReg-2 (SEQ ID Nos. 3 and 4), displays up-regulatory properties. 3-PGA activation studies of the partially purified enzyme (specific activity of 9.1 μmol/min/mg) yielded an $A_{0.5}$ of 0.035 mM, an 11.8- and 4.6-fold enhancement over the parental revertant R32 and WT. Although the increase in affinity for 3-PGA was less pronounced for UpReg-2 than for UpReg-1, the response to Pi was more dramatic for UpReg-2 than for UpReg-1 . Pi inhibition studies in the presence of 0.1 and 0.001 mM 3-PGA revealed $I_{0.5}$ constants of 6.2 and 1.3 mM, respectively. Hence, UpReg-2 shows a higher resistance to Pi inhibition than UpReg-1.

EXAMPLE 3: Generation and Identification of Up-Regulatory ADPG-PP by N-Terminal Deletions of the Large Subunit

A DNA fragment encoding a truncated N-terminal peptide missing the first 17 amino acids of the large subunit was amplified using PCR with the primers shown in FIGS. 5I and 5J. The PCR product was purified, digested with NcoI and NheI, subcloned into pML7 (Fig. 1) and expressed in the *E. coli* strain JM101 (Messing, *Meth. Enzymol.* **101**:20-78, 1983) containing pML10 (Fig. 2).

The resulting truncated large subunit was co-expressed with the wild-type small subunit in *E. coli* to form the ΔN17-LS enzyme. The small subunit was expressed at 109% and the large subunit at 114% of the wild-type level, as quantified by ELISA. The ΔN17-LS enzyme was partially purified to a final specific activity of 20 units/mg under standard assay conditions

and estimated to be approximately 51% pure based on video image analysis of a purified protein fraction on Commassie Brilliant Blue-stained polyacrylamide gels. Kinetic analysis of the partially purified ΔN17-LS enzyme showed no effect on the binding constants for the substrates ATP and glucose-1-P, while the enzyme's affinity for the co-factor Mg^{2+} increased slightly (Fig. 5). A marked change in allosteric regulatory behavior of the ΔN17-LS enzyme was observed. The ΔN17-LS enzyme required only 0.015 mM 3-PGA for 50% activation, reflecting a ten-fold greater sensitivity to 3-PGA than the wild-type enzyme (Fig. 5). The ΔN17-LS enzyme was also less sensitive to inhibition by Pi (Fig. 6). In the presence of 0.25 mM 3-PGA, the ΔN17-LS enzyme had an $I_{0.5}$ of 2.0 mM (Fig. 5), a value more than 28-fold greater than that exhibited by the wild-type enzyme (Ballicora *et al.*, *Plant Physiol.* **109**:245-251, 1995). Even at 0.025 mM 3-PGA, 0.13 mM Pi was required to inhibit the ΔN-L17 enzyme activity by 50%. Comparing the calculated $I_{0.5}$s of the native and recombinant enzymes, this decrease in sensitivity to Pi is dramatic. In contrast to the shallow linear increases in the ratio of $I_{0.5}$/3-PGA exhibited by the native and recombinant wild-type enzymes, the ΔN17-LS enzyme showed an increasing resistance to Pi inhibition as 3-PGA levels increase (Fig. 4). Thus, this 17 amino-acid N-terminal region of the large subunit is essential for proper allosteric regulation, since its removal increases the enzyme's sensitivity to the allosteric activator 3-PGA 10-fold and decreases its sensitivity to the allosteric inhibitor Pi 5- to 16-fold.

EXAMPLE 4: Transfer and Expression of Regulatory Mutants and Wildtype Potato Adp-glucose Pyrophosphorylase in Arabidopsis

Expression of up-regulatory allosteric large subunits of ADPG-PP in photosynthetic tissues of

Arabidopsis increases vegetative growth rate and seed yield.

The allosteric regulatory properties of the ADPG-PP enzymes formed from the mutant large-subunit sequences together with normal small-subunit sequences are shown in Fig. 7. Sensitivity to the activator 3-PGA ranges from 0.002 mM (upReg1), 0.14 mM (wild-type: "W.T."), to 4 mM (M345). Since the enzyme is also inhibited by Pi, a more accurate view of the differences in the allosteric regulation of the mutant enzymes is the $I_{0.5}$ / [3-PGA] ratio, where $I_{0.5}$ is the amount of Pi required to produce 50% inhibition of enzyme activity ($I_{0.5}$) at a known concentration of 3-PGA. $I_{0.5}$ / [3-PGA] ratios range from a value of 24 for UpReg1 to 1.4 for the W.T. enzyme to 0.6 for the down-regulatory mutant M345. Specifically, the up-regulatory type enzymes have a higher affinity for the activator 3-phosphoglyceric acid (3-PGA) and/or higher resistance to the inhibitor inorganic phosphate (Pi) than the W.T. enzyme. Expression of these up-regulatory ADPG-PP enzymes results in an increased production of glycogen (starch-like) in *Escherichia coli*, and it is expected that expression of these gene sequences in higher plants would increase starch biosynthesis and levels in cells capable of starch accumulation. The effect of this increased starch production on plant growth and development will depend on the organs of the plant in which these up-regulated enzymes are expressed. In non-photosynthetic storage organs such as tubers, developing seeds, fruits and roots, expression of unregulated or up-regulated forms of ADPG-PP would not be expected to have a serious effect on plant growth and development because starch serves as a primary form of fixed carbon.

In contrast, expression of up-regulated ADPG-PP in leaves may have deleterious consequences on plant growth and development because the allosteric regulatory properties of ADPG-PP are believed to be one of several

important processes that are responsible for controlling the distribution of fixed carbon into sucrose and starch in leaves and other photosynthetic competent tissues. One hypothesis (Eichelmann and Laisk, *Plant Physiol.* **106**:679-687, 1994) is that the bulk of the fixed carbon is converted into sucrose and that starch is made only when the rate of sucrose synthesis is saturated. Under conditions where sucrose synthesis is saturated, the ratio of activator to inhibitor (3-PGA/Pi) is high enough to activate ADPG-PP, thereby allowing starch synthesis to occur. In the dark, the 3-PGA/Pi ratio is low, which suppresses the enzyme activity of ADPG-PP and, in turn, limits starch synthesis. The normal allosteric regulatory properties of ADPG-PP account for the diurnal oscillation of starch synthesis during the day and net breakdown of starch at night. Based on this hypothesis regarding the critical role of ADPG-PP in leaf starch metabolism, it is expected that the expression of unregulated and up-regulated forms of ADPG-PP in leaf cells would drastically alter the normal partitioning of fixed carbon between sucrose and starch and cause much higher starch synthesis to occur. These events, in turn, would reduce sucrose availability to the remainder of the plant which could disrupt normal plant growth and development processes. Stark et al. (*Science* **258**:287-292, 1992) showed that expression of the allosteric unregulated glgC16 gene by the constitutive CaMV35S promoter produced a transgenic potato plant that could not survive without sucrose.

More recent evidence indicates that starch does not merely serve as a transient reserve to support the metabolic activities of the plant during the night, but also as a transient sink to accommodate excess photosynthate. Ludewig et al. (*FEBS Lett.* **429**: 147-151) have shown that there is a direct correlation between the capacity of starch synthesis and the rate of photosynthesis at elevated CO_2. Likewise, recent data

from J. Sun, G.E. Edwards, and T.W. Okita (unpublished) exhibited a significant correlation between the rates of starch synthesis and CO_2 assimilation, and between the rates of starch synthesis and accumulative leaf area. These results indicate that leaf starch plays an important role as a transient "sink" for excess carbon formed during photosynthesis and thereby alleviates any potential feedback of photosynthesis. If this view is correct, increased starch production in leaf tissue should lead to increased photosynthesis and, in turn, higher productivity. This was tested by transforming *Arabidopsis* plants with the various up-regulated potato large subunit sequences shown in Fig. 5 under the control of the *Arabidopsis* ribulose bisphosphate carboxylase small-subunit promoter.

EXAMPLE 5: Construction of Plant Expression Plasmids Containing The Wild-Type and Mutant ADPG-PP Large Subunits

The *Arabidopsis* ribulose bisphosphate small-subunit (ats1A) promoter (Krebbers et al., *Plant Mol. Biol.* 11:745-759, 1989)) and transit leader coding sequences were amplified using synthetic primers using Taq DNA polymerase. The 5' primer contained a XhoI-XbaI-BamHI sequence 5' (underlined) to the ats1A nucleotide sequence beginning with base -1701 from the translation start of the ats1A gene. The 3' primer had a NcoI-SacI sequence (underlined) at the complementary nucleotide located at +164 from the translation start. The amplified 1897-nucleotide DNA fragment was ethanol precipitated, resuspended in TE, and digested with XhoI and SacI. The digested DNA was then resolved by agarose gel electrophoresis, and then purified by binding and elution on DEAE-membrane filters. The purified DNA fragment was then collected by ethanol precipitation and

then cloned into the XhoI and SacI sites of pBluescript II to give pHI-10 (Fig. 6):

5' primer GCTCGAGTCTAGAGGATCCGTGGTCGAGATTGTGTATTATTCTTTAG

3' primer CGAGCTCGCCATGGCAGTTAACTCTTCCGCCGTTGCTTG

To form a gene fusion between the ats1A promoter and transit leader sequences to the potato ADPG-PP large-subunit sequences, the coding sequences from UpReg1, R4, R20, R32, M27, M345, and wild-type large-subunit sequences were removed from the plasmid DNA (Iglesias et al., *J. Biol. Chem.* **268**:1081-1086, 1993) by digestion with NcoI and SacI and the resulting DNA fragment cloned into the relevant restriction sites of pHI-10 to give pHI-11 to pHI-17 (Fig. 11).

The ats1A-potato large subunit cassettes contained within a XbaI/SacI DNA fragment were then cloned into the XbaI and SacI sites of the T-DNA binary vector pHI-32, a derivative of pIG-121 (Ohta et al., *Plant Cell Physiol.* **31**:805-813, 1990) to give pHI-33 to pHI-39 (Fig. 9).

EXAMPLE 6: Transformation into Arabidopsis TL46

The pHI-33 to -39 series of plasmid DNAs were transferred into *Agrobacterium* GV3101 using standard methods (An, *Methods Enzymol.* **153**:292-305, 1987). These *Agrobacterium* lines were then used to transformed *Arabidopsis* Columbia line TL-46. TL-46 is a starch-deficient line (Lin et al., *Plant Physiol.* **99**:1175-1181, 1988) which is defective for the leaf ADPG-PP large subunit. Transformation was accomplished by co-cultivation of leaf sections (discs) of *Arabidopsis* TL-46, selection of kanamycin- and hygromycin-resistant calli and the regeneration of plants by methods described by van der Graaf and Hooykaas (*Plant Cell*

Reports **15**:572-577, 1966). Alternatively, transgenic *Arabidopsis* plants containing pHI-33 to pHI-39 were obtained by the vacuum infiltration method (Bechtold et al., *C.R. Acad. Sci. Paris* **316**:1194-1199, 1993).

Third generation progeny of the transgenic *Arabidopsis* plants containing pHI-33 to pHI-39 were evaluated for their growth properties. When germinated and cultured on MS media with or without 2% sucrose, transgenic plants expressing up-regulatory AGPase (UpReg1 and R4) and wildtype (R20 and WT) type AGPases grew considerably faster than control plants (see Fig. 10 and Fig. 13). Moreover, when transferred to soil, several of the transgenic plants bearing UpReg-1 and R4 produced larger quantities of seed at the end of 8 weeks of growth than normal plants (Fig. 8).

The observed phenotypic traits expressed by these transgenic plants are striking (Figs. 8 and 13). They include higher rates of development and growth, and variation in leaf size, seed weights, and seed yields. None of these phenotypes are strictly inter-related although there appears to be several general trends which can be summarized as follows:

1. Most of the plants containing the up-regulatory (UpReg1 and R4) to wild-type (R20 and W.T.) potato AGPase LS grew faster than normal Columbia plants (Fig. 13). In contrast, plants containing the down-regulatory AGPase LS genes, R32 and M345, grew at the same rate as Columbia plants suggesting that increased growth rates are due to an increase in enzyme activity levels in these transgenic plants.

2. Increased seed yield but not seed weight is correlated with the presence of up-regulatory allosteric potato AGPase UpReg1 and R4 (Fig.

8). This observation suggests that the increase in seed yield is due to increased enzyme activity via the up-regulatory response.

3. There is little correlation between increased leaf size and allosteric regulatory behavior and, in turn, enzyme activity of AGPase (Fig. 13). Large leaves are not only evident in the up-regulatory plants such as UpReg1 but also in the down-regulatory plants such as R32 and M345. Alternatively, the up-regulatory plant, R4, which grows faster than normal Columbia plants, has smaller leaves. These findings suggest that the variation in leaf size is not due to enzyme function but due to some other non-catalytic interaction by the transgene AGPase LS.

4. There is little correlation between larger seed weight and allosteric regulatory behavior and, in turn, enzyme activity of AGPase (Fig. 13). Larger seed weights are observed in plants containing either up-regulatory or down regulatory AGPase LS genes. Larger seed weights, however, are correlated with a reduction in seed yield. The reduction in seed yield is due to the reduced number of seeds per pod.

Having illustrated and described the principles of the present invention, it should be apparent to persons skilled in the art that the invention can be modified in arrangement and detail without departing from such principles. The invention encompasses all modifications that are within the spirit and scope of the appended claims.

WHAT IS CLAIMED IS:

1. A method for producing a nucleic acid encoding an up-regulated mutant ADPG-PP enzyme, comprising the steps of:

(i) providing a nucleic acid that encodes a large subunit of an allosterically regulated ADPG-PP enzyme, the ADPG-PP enzyme consisting of the large subunit and a small subunit;

(ii) introducing a first mutation to said nucleic acid to produce a first mutant nucleotide;

(iii) identifying a first mutant nucleic acid, a protein product of which, when co-expressed in a cell with a wild-type ADPG-PP small subunit, is defective in allosteric regulation;

(iv) introducing a second mutation to said nucleic acid to produce a second mutant nucleic acid;

(v) identifying a second mutant nucleic acid, a protein product of which, when co-expressed in a cell with a wild-type ADPG-PP small subunit, at least partially restores allosteric regulation;

(vi) sequencing the second mutant nucleic acid, the protein product of which at least partially restores allosteric regulation so as to characterize the first and second mutations;

(vii) producing a third mutant nucleic acid having the second, but not the first mutation; and

(viii) identifying a third mutant nucleic acid, a protein product of which has greater biological activity than a wild-type ADPG-PP enzyme, thereby producing and identifying a nucleic acid encoding an up-regulated mutant ADPG-PP enzyme.

2. The method of claim 1, wherein the nucleic acid that encodes the large subunit of ADPG-PP is native to a plant.

3. A nucleic acid molecule produced by the method of claim 1.

4. A peptide encoded by the nucleic acid molecule of claim 3.

5. A cell comprising the nucleic acid molecule of claim 3.

6. A method for producing a nucleic acid encoding an up-regulated mutant ADPG-PP enzyme, comprising the steps of:

(i) providing a nucleic acid that encodes a large subunit of an allosterically regulated ADPG-PP enzyme, the ADPG-PP enzyme consisting of the large subunit and a small subunit;

(ii) introducing a first mutation to said nucleic acid to produce a first mutant nucleic acid;

(iii) cloning said first mutant nucleic acid into a low copy-number vector;

(iv) transforming said cloned first mutant nucleic acid into a population of cells defective in ADPG-PP function;

(v) expressing said cloned first mutant nucleic acid in said cells; and

(vi) screening said transformed cells to identify a nucleic acid coding for an up-regulated mutant ADPG-PP enzyme, thereby producing and identifying a nucleic acid encoding an up-regulated mutant ADPG-PP enzyme.

7. The method of claim 6, wherein the nucleic acid that encodes the large subunit of ADPG-PP is native to a plant.

8. A nucleic acid molecule produced by the method of claim 6.

9. A peptide encoded by the nucleic acid molecule of claim 8.

10. A cell comprising the nucleic acid molecule of claim 8.

11. A method for producing a nucleic acid encoding an up-regulated mutant ADPG-PP enzyme, comprising the steps of:

(i) providing a nucleic acid that encodes a large subunit of an allosterically regulated ADPG-PP enzyme, the ADPG-PP enzyme consisting of the large subunit and a small subunit;

(ii) deleting a terminal portion of said of the nucleic acid molecule to produce a first mutant nucleic acid;

(iii) cloning said first mutant nucleic acid into a low copy-number vector;

(iv) transforming said cloned first mutant nucleic acid into a population of cells defective in ADPG-PP function;

(v) expressing said cloned first mutant nucleic acid in said cells; and

(vi) screening said cells to identify a nucleic acid encoding an up-regulated mutant ADPG-PP enzyme, thereby producing and identifying a nucleic acid encoding an up-regulated mutant ADPG-PP enzyme.

12. The method of claim 11 wherein the nucleic acid that encodes the large subunit of ADPG-PP is native to a plant.

13. A nucleic acid molecule produced by the method of claim 11

14. A peptide encoded by the nucleic acid molecule of claim 13.

15. A cell comprising the nucleic acid molecule of claim 13.

16. A nucleic acid molecule, comprising a nucleic acid sequence selected from the group consisting of the sequences shown in SEQ. I.D. Nos. 1, 3, 5, 7, 9, and 11.

17. A polypeptide, comprising an amino acid sequence selected from the group consisting of the sequences shown in SEQ. I.D. Nos. 2, 4, 6, 8, 10, and 12.

18. A cell, comprising the nucleic acid molecule of claim 16.

19. The cell of claim 18, selected from a group consisting of a plant cell, a bacterial cell, and a fungal cell.

20. The cell of claim 19, exhibiting at least one characteristic selected from the group consisting of increased yield compared to a wild-type plant cell, increased productivity compared to a wild-type plant cell, increased starch production compared to a wild-type plant cell, increased size compared to a wild-type plant cell, increased rate of growth compared to a wild-type plant cell, and increased number of seeds compared to a wild-type plant cell.

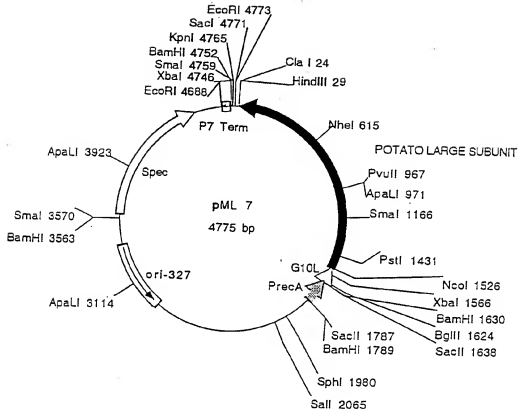

Fig. 1

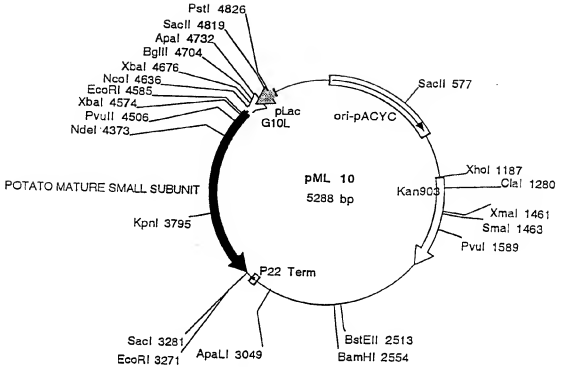

Fig. 2

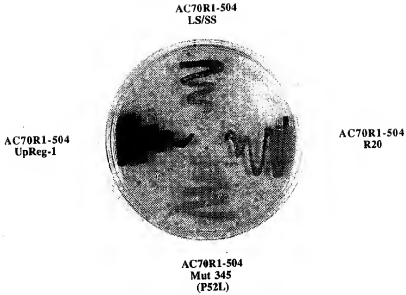

Fig. 3

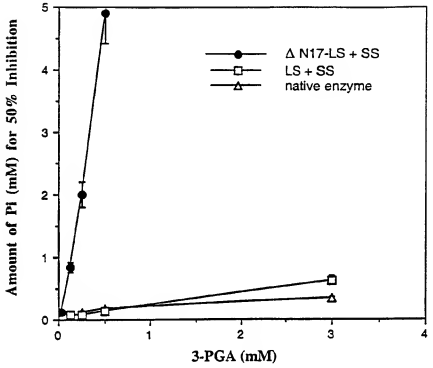

Fig. 4

	ATP	glucose 1-phosphate (mM)	Mg2+	3-PGA
native[1]	0.19	0.14	-	0.4
wildtype[2]	0.12	0.04	2.00	0.160
mutant 345[3]	0.34	0.30	2.57	4.50
R20	0.17 ± .001	0.08 ± .026	2.70 ± 0.60	0.117 ± .036
UpiReg-1	0.10 ± .003	0.11 ± .030	2.24 ± 0.50	0.002 ± 0.0001
ΔN17-LS	0.16 ± 0.01	0.11 ± 0.01	1.2 ± 0.1	0.015 ± 0.001

Substrate binding ($K_{0.5}$) and 3-PGA activator ($A_{0.5}$) affinity properties of various ADP-glucose pyrophosphorylases. [1]Sowokinos and Preiss, 1982 [2]Ballicora *et al.*, 1995 [3]Greene *et al.*, 1996

Fig. 5

3-PGA (mM)	$I_{0.5}$ (mM)				
	wildtype[1] enzyme	mutant 345[2] enzyme	R20	UpReg-1	ΔN17-LS
0.025	n.d.[3]	n.d.	n.d.	0.60	0.13
0.125	n.d.	n.d.	0.21 ± .03	3.1	0.84
0.25	0.07	n.d.	0.57 ± .09	4.7	2.0
0.50	0.15	n.d.	0.70 ± .20	n.d.	n.d.
1.0	n.d.	0.68	n.d.	n.d.	n.d.
2.25	n.d.	0.84	n.d.	n.d.	n.d.
3.0	0.63	13.00	n. d.	n.d.	n.d.

Pi inhibition constants ($I_{0.5}$) of various ADP-glucose pyrophosphorylases. [1]Ballicora *et al.*, 1995 [2]Greene *et al.*, 1996 [3]n.d., not determined

Fig. 6

AGP Line	$A_{0.5}$ (mM)	$I_{0.5}$ (mM)	$I_{0.5}$ / [3-PGA]
UpReg1	0.002	3.1	24.0
R4	0.09	1.2	4.8
R20	0.12	0.21	1.7
W.T.	0.14	0.17	1.4
R32	0.41	0.32	1.3
M345 (P52L)	4.0	2.5	0.6

$A_{0.5}$ is the amount of 3-PGA required to give 50% activation.
$I_{0.5}$ is the amount of Pi required to inhibit the enzyme 50% in the presence of a known amount of 3-PGA.

Fig. 7

AGP gene	line	seeds: Number of seeds (+/- 5%)	seeds: properties	relative number compared to Columbia WT
UpReg1	1	22,000	white	0.88
	2	43,000	normal	1.72
	3	35,000	normal	1.40
	4	3,000	large	0.12
R4	1	45,000	normal	1.80
	3	39,000	normal	1.56
Columbia WT		25,000	normal	1.0

Seed yields from eight greenhouse grown plants of each line (T3 generation)

Fig. 8

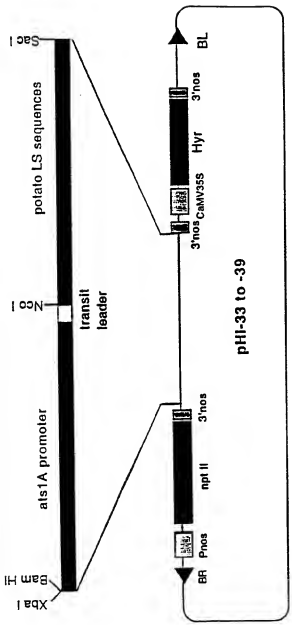

Fig. 9

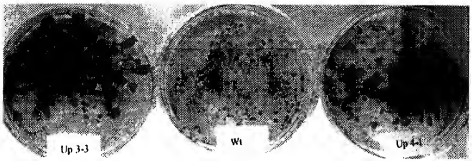

Fig. 10

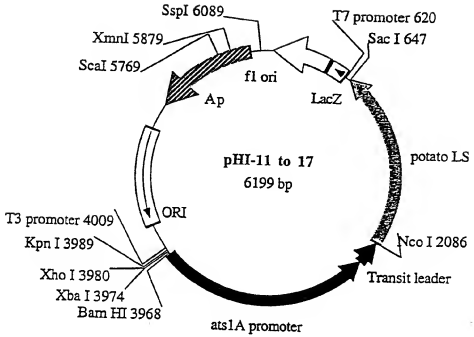

Fig. 11

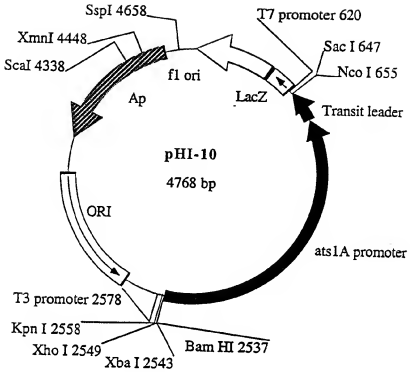

Fig. 12

AGP gene	Line	Growth rate as compared to wild type (early stage)	Seed weight (mg/seed)	pods *1 (cm)	leaves *2 (cm) (late stage)	leaf phenotype as compared to wild type
Uplteg-1	1	faster	0.031	0.80-0.96	3.20/1.92, 4.16/1.92, 3.52/1.60	large
	2	faster	0.022	1.12-1.28	3.52/1.60, 3.20/1.60, 3.52/1.60	large
	3	faster	0.027	0.96-1.12	3.36/1.92, 3.84/2.08, 3.36/1.76	large
	4	faster	0.043	0.64-0.96	2.88/1.60, 3.20/1.60, 3.20/1.92	large
R4	1		0.023			
	3	faster	0.022	0.96-1.12	2.88/1.44, 2.88/1.44, 2.88/1.28	small
R20	1	faster	0.023	0.96-1.28	3.20/1.76, 3.20/1.60, 3.20/1.44	large
	2	slower	0.034	0.32-0.96	5.44/2.24, 5.12/3.84, 3.84/2.24	very large
	3	faster	0.024	0.96-1.12	3.84/1.92, 3.84/1.92, 3.52/1.76	large
rWT	1	faster	0.036	0.64-1.12	3.84/2.24, 3.84/1.92, 3.68/1.92	large
	2	same	0.028	1.12-1.44	3.84/1.60, 4.48/1.60, 4.80/1.92	large
	3	faster	0.038	0.80-0.96	3.52/1.92, 3.52/2.08, 3.52/2.24	large
R32	2	same	0.039	0.64-0.96	3.84/2.88, 3.20/2.24, 3.20/1.92	large
M345	1	same	0.022	0.96-1.28	3.20/1.28, 2.88/1.60, 3.20/1.60	same
	2	same	0.022	0.96-1.28	2.88/1.92, 2.56/1.92, 2.56/1.60	same
	3	same	0.043	0.80-1.12	4.80/3.20, 3.84/2.56, 4.16/1.92	very large
TL46		somewhat slower	0.025	0.96-1.44	3.52/1.60, 3.20/1.28, 3.52/1.60	same
Wild type			0.027	1.28-1.44	3.20/1.28, 3.52/1.28, 2.72/1.60	

Phenotypic Properties of Transgenic Plants (T2 Generation)
*1 Twenty pods were used to measure the lengths.
*2 Three rosetta leaves were used to measure the widths and lengths.

Fig. 13

SEQUENCE LISTING

<110> Okita, Thomas et al

<120> Regulatory Mutants of ADP-Glucose Pyrophosphorylase and Related Compostions and Methods

<130> 47759

<140>
<141>

<150> 60/055,410
<151> 1997-08-07

<160> 22

<170> PatentIn Ver. 2.0

<210> 1
<211> 1392
<212> DNA
<213> Solanum tuberosum

<220>
<221> CDS
<222> (1)..(1392)

<400> 1

```
atg gcc tct gtg atc act act gaa aat gac aca cag act gtg ttc gta   48
Met Ala Ser Val Ile Thr Thr Glu Asn Asp Thr Gln Thr Val Phe Val
  1               5                  10                  15

gat atg cca cgt ctt gag aga cgc cgg gca aat cca aag gat gtg gct   96
Asp Met Pro Arg Leu Glu Arg Arg Arg Ala Asn Pro Lys Asp Val Ala
            20                  25                  30

gca gtc ata ctg gga gga gga aaa ggg acc aag tta ttc cca ctt aca   144
Ala Val Ile Leu Gly Gly Gly Lys Gly Thr Lys Leu Phe Pro Leu Thr
        35                  40                  45

agt aga act gca acc cct gct gtt ccg gtt gga gga tgc tac agg cta   192
Ser Arg Thr Ala Thr Pro Ala Val Pro Val Gly Gly Cys Tyr Arg Leu
    50                  55                  60

ata gac atc cca atg agc aac tgt atc aac agt gct att aac aag att   240
Ile Asp Ile Pro Met Ser Asn Cys Ile Asn Ser Ala Ile Asn Lys Ile
65                  70                  75                  80
```

```
ttt gtg ctg aca cag tac aat tct gct ccc ctg aat cgt cac att gct   288
Phe Val Leu Thr Gln Tyr Asn Ser Ala Pro Leu Asn Arg His Ile Ala
                85                  90                  95

cga aca tat ttt ggc aat ggt gtg agc ttt gga gat gga ttt gtc gag   336
Arg Thr Tyr Phe Gly Asn Gly Val Ser Phe Gly Asp Gly Phe Val Glu
            100                 105                 110

gta cta gct gca act cag aca ccc ggg gaa gca gga aaa aaa tgg ttt   384
Val Leu Ala Ala Thr Gln Thr Pro Gly Glu Ala Gly Lys Lys Trp Phe
        115                 120                 125

caa gga aca gca gat gct gtt aga aaa ttt ata tgg gtt ttt gag gac   432
Gln Gly Thr Ala Asp Ala Val Arg Lys Phe Ile Trp Val Phe Glu Asp
    130                 135                 140

gct aag aac aag aat att gaa aat atc gtt gta cta tct ggg gat cat   480
Ala Lys Asn Lys Asn Ile Glu Asn Ile Val Val Leu Ser Gly Asp His
145                 150                 155                 160

ctt tat agg atg gat tat atg gag ttg gtg cag aac cat att gac agg   528
Leu Tyr Arg Met Asp Tyr Met Glu Leu Val Gln Asn His Ile Asp Arg
                165                 170                 175

aat gct gat att act ctt tca tgt gca cca gct gag gac agc cga gca   576
Asn Ala Asp Ile Thr Leu Ser Cys Ala Pro Ala Glu Asp Ser Arg Ala
            180                 185                 190

tca gat ttt ggg ctg gtc aag att gac agc aga ggc aga gta gtc cag   624
Ser Asp Phe Gly Leu Val Lys Ile Asp Ser Arg Gly Arg Val Val Gln
        195                 200                 205

ttt gct gaa aaa cca aaa ggt ttt gat ctt aaa gca atg caa gta gat   672
Phe Ala Glu Lys Pro Lys Gly Phe Asp Leu Lys Ala Met Gln Val Asp
    210                 215                 220

act act ctt gtt gga tta tct cca caa gat gcg aag aaa tcc ccc tat   720
Thr Thr Leu Val Gly Leu Ser Pro Gln Asp Ala Lys Lys Ser Pro Tyr
225                 230                 235                 240

att gct tca atg gga gtt tat gta ttc aag aca gat gta ttg ttg aag   768
Ile Ala Ser Met Gly Val Tyr Val Phe Lys Thr Asp Val Leu Leu Lys
                245                 250                 255

ctc ttg aaa tgg agc tat ccc act tct aat gat ttt ggc tct gaa att   816
Leu Leu Lys Trp Ser Tyr Pro Thr Ser Asn Asp Phe Gly Ser Glu Ile
            260                 265                 270
```

```
ata cca gca gct att gac gat tac aat gtc caa gca tac att ttc aaa   864
Ile Pro Ala Ala Ile Asp Asp Tyr Asn Val Gln Ala Tyr Ile Phe Lys
        275                 280                 285

gac tat tgg gaa gac att gga aca att aaa tcg ttt tat aat gct agc   912
Asp Tyr Trp Glu Asp Ile Gly Thr Ile Lys Ser Phe Tyr Asn Ala Ser
    290                 295                 300

ttg gca ctc aca caa gag ttt cca gag ttc caa ttt tac gat cca aaa   960
Leu Ala Leu Thr Gln Glu Phe Pro Glu Phe Gln Phe Tyr Asp Pro Lys
305                 310                 315                 320

aca cct ttt tac aca tct cct agg ttc ctt cca cca acc aag ata gac   1008
Thr Pro Phe Tyr Thr Ser Pro Arg Phe Leu Pro Pro Thr Lys Ile Asp
                325                 330                 335

aat tgc aag att aag gat gcc ata atc tct cat gga tgt ttc ttg cga   1056
Asn Cys Lys Ile Lys Asp Ala Ile Ile Ser His Gly Cys Phe Leu Arg
            340                 345                 350

gat tgt tct gtg gaa cac tcc ata gtg ggt gaa aga tcg cgc tta gat   1104
Asp Cys Ser Val Glu His Ser Ile Val Gly Glu Arg Ser Arg Leu Asp
        355                 360                 365

tgt ggt gtt gaa ctg aag gat act ttc atg atg gga gca gac tac tac   1152
Cys Gly Val Glu Leu Lys Asp Thr Phe Met Met Gly Ala Asp Tyr Tyr
    370                 375                 380

caa aca gaa tct gag att gcc tcc ctg tta gca gag ggg aaa gta ccg   1200
Gln Thr Glu Ser Glu Ile Ala Ser Leu Leu Ala Glu Gly Lys Val Pro
385                 390                 395                 400

att gga att ggg gaa aat aca aaa ata agg aaa tgt atc att gac aag   1248
Ile Gly Ile Gly Glu Asn Thr Lys Ile Arg Lys Cys Ile Ile Asp Lys
                405                 410                 415

aac gca aag ata gga aag aat gtt tca atc ata aat aaa gac ggt gtt   1296
Asn Ala Lys Ile Gly Lys Asn Val Ser Ile Ile Asn Lys Asp Gly Val
            420                 425                 430

caa gag gca gac cga cca gag gaa gga ttc tac ata cga tca ggg ata   1344
Gln Glu Ala Asp Arg Pro Glu Glu Gly Phe Tyr Ile Arg Ser Gly Ile
        435                 440                 445

atc att ata tta gag aaa gcc aca att aga gat gga aca gtc ata tga   1392
Ile Ile Ile Leu Glu Lys Ala Thr Ile Arg Asp Gly Thr Val Ile
    450                 455                 460
```

<210> 2
<211> 463
<212> PRT
<213> Solanum tuberosum

<400> 2
Met Ala Ser Val Ile Thr Thr Glu Asn Asp Thr Gln Thr Val Phe Val
1 5 10 15

Asp Met Pro Arg Leu Glu Arg Arg Arg Ala Asn Pro Lys Asp Val Ala
20 25 30

Ala Val Ile Leu Gly Gly Gly Lys Gly Thr Lys Leu Phe Pro Leu Thr
35 40 45

Ser Arg Thr Ala Thr Pro Ala Val Pro Val Gly Gly Cys Tyr Arg Leu
50 55 60

Ile Asp Ile Pro Met Ser Asn Cys Ile Asn Ser Ala Ile Asn Lys Ile
65 70 75 80

Phe Val Leu Thr Gln Tyr Asn Ser Ala Pro Leu Asn Arg His Ile Ala
85 90 95

Arg Thr Tyr Phe Gly Asn Gly Val Ser Phe Gly Asp Gly Phe Val Glu
100 105 110

Val Leu Ala Ala Thr Gln Thr Pro Gly Glu Ala Gly Lys Lys Trp Phe
115 120 125

Gln Gly Thr Ala Asp Ala Val Arg Lys Phe Ile Trp Val Phe Glu Asp
130 135 140

Ala Lys Asn Lys Asn Ile Glu Asn Ile Val Val Leu Ser Gly Asp His
145 150 155 160

Leu Tyr Arg Met Asp Tyr Met Glu Leu Val Gln Asn His Ile Asp Arg
165 170 175

Asn Ala Asp Ile Thr Leu Ser Cys Ala Pro Ala Glu Asp Ser Arg Ala
180 185 190

Ser Asp Phe Gly Leu Val Lys Ile Asp Ser Arg Gly Arg Val Val Gln
195 200 205

Phe Ala Glu Lys Pro Lys Gly Phe Asp Leu Lys Ala Met Gln Val Asp

```
    210                 215                 220

Thr Thr Leu Val Gly Leu Ser Pro Gln Asp Ala Lys Lys Ser Pro Tyr
225                 230                 235                 240

Ile Ala Ser Met Gly Val Tyr Val Phe Lys Thr Asp Val Leu Leu Lys
                245                 250                 255

Leu Leu Lys Trp Ser Tyr Pro Thr Ser Asn Asp Phe Gly Ser Glu Ile
            260                 265                 270

Ile Pro Ala Ala Ile Asp Asp Tyr Asn Val Gln Ala Tyr Ile Phe Lys
        275                 280                 285

Asp Tyr Trp Glu Asp Ile Gly Thr Ile Lys Ser Phe Tyr Asn Ala Ser
    290                 295                 300

Leu Ala Leu Thr Gln Glu Phe Pro Glu Phe Gln Phe Tyr Asp Pro Lys
305                 310                 315                 320

Thr Pro Phe Tyr Thr Ser Pro Arg Phe Leu Pro Pro Thr Lys Ile Asp
                325                 330                 335

Asn Cys Lys Ile Lys Asp Ala Ile Ile Ser His Gly Cys Phe Leu Arg
            340                 345                 350

Asp Cys Ser Val Glu His Ser Ile Val Gly Glu Arg Ser Arg Leu Asp
        355                 360                 365

Cys Gly Val Glu Leu Lys Asp Thr Phe Met Met Gly Ala Asp Tyr Tyr
    370                 375                 380

Gln Thr Glu Ser Glu Ile Ala Ser Leu Leu Ala Glu Gly Lys Val Pro
385                 390                 395                 400

Ile Gly Ile Gly Glu Asn Thr Lys Ile Arg Lys Cys Ile Ile Asp Lys
                405                 410                 415

Asn Ala Lys Ile Gly Lys Asn Val Ser Ile Ile Asn Lys Asp Gly Val
            420                 425                 430

Gln Glu Ala Asp Arg Pro Glu Glu Gly Phe Tyr Ile Arg Ser Gly Ile
        435                 440                 445

Ile Ile Ile Leu Glu Lys Ala Thr Ile Arg Asp Gly Thr Val Ile
    450                 455                 460
```

<210> 3
<211> 1392
<212> DNA
<213> Solanum tuberosum

<220>
<221> CDS
<222> (1)..(1392)

<400> 3
atg gcc tct gtg atc act act gaa aat gac aca cag act gtg ttc gta 48
Met Ala Ser Val Ile Thr Thr Glu Asn Asp Thr Gln Thr Val Phe Val
1 5 10 15

gat atg cca cgt ctt gag aga cgc cgg gca aat cca aag gat gtg gct 96
Asp Met Pro Arg Leu Glu Arg Arg Arg Ala Asn Pro Lys Asp Val Ala
20 25 30

gca gtc ata ctg gga gga gga gaa ggg acc aag tta ttc cca ctt aca 144
Ala Val Ile Leu Gly Gly Gly Glu Gly Thr Lys Leu Phe Pro Leu Thr
35 40 45

agt aga act gca acc cct gct gtt ccg gtt gga gga tgc tac agg cta 192
Ser Arg Thr Ala Thr Pro Ala Val Pro Val Gly Gly Cys Tyr Arg Leu
50 55 60

ata gac atc cca atg agc aac tgt atc aac agt gct att aac aag att 240
Ile Asp Ile Pro Met Ser Asn Cys Ile Asn Ser Ala Ile Asn Lys Ile
65 70 75 80

ttt gtg ctg aca cag tac aat tct gct ccc ctg aat cgt cac att gct 288
Phe Val Leu Thr Gln Tyr Asn Ser Ala Pro Leu Asn Arg His Ile Ala
85 90 95

cga aca tat ttt ggc aat aat gtg agc ttt gga gat gga ttt gtc gag 336
Arg Thr Tyr Phe Gly Asn Asn Val Ser Phe Gly Asp Gly Phe Val Glu
100 105 110

gta cta gct gca act cag aca ccc ggg gaa gca gga aaa aaa tgg ttt 384
Val Leu Ala Ala Thr Gln Thr Pro Gly Glu Ala Gly Lys Lys Trp Phe
115 120 125

caa gga aca gca gat gct gtt aga aaa ttt ata tgg gtt ttt gag gac 432
Gln Gly Thr Ala Asp Ala Val Arg Lys Phe Ile Trp Val Phe Glu Asp
130 135 140

gct aag aac aag aat att gaa aat atc gtt gta cta tct ggg gat cat 480
Ala Lys Asn Lys Asn Ile Glu Asn Ile Val Val Leu Ser Gly Asp His

```
145                 150                 155                 160

ctt tat agg atg gat tat atg gag ttg gtg cag aac cat att gac agg   528
Leu Tyr Arg Met Asp Tyr Met Glu Leu Val Gln Asn His Ile Asp Arg
                165                 170                 175

aat gct gat att act ctt tca tgt gca cca gct gag gac agc cga gca   576
Asn Ala Asp Ile Thr Leu Ser Cys Ala Pro Ala Glu Asp Ser Arg Ala
            180                 185                 190

tca gat ttt ggg ctg gtc aag att gac agc aga ggc aga gta gtc cag   624
Ser Asp Phe Gly Leu Val Lys Ile Asp Ser Arg Gly Arg Val Val Gln
        195                 200                 205

ttt gct gaa aaa cca aaa ggt ttt gat ctt aaa gca atg caa gta gat   672
Phe Ala Glu Lys Pro Lys Gly Phe Asp Leu Lys Ala Met Gln Val Asp
    210                 215                 220

act act ctt gtt gga tta tct cca caa gat gcg aag aaa tcc ccc tat   720
Thr Thr Leu Val Gly Leu Ser Pro Gln Asp Ala Lys Lys Ser Pro Tyr
225                 230                 235                 240

att gct tca atg gga gtt tat gta ttc aag aca gat gta ttg ttg aag   768
Ile Ala Ser Met Gly Val Tyr Val Phe Lys Thr Asp Val Leu Leu Lys
                245                 250                 255

ctc ttg aaa tgg agc tat ccc act tct aat gat ttt ggc tct gaa att   816
Leu Leu Lys Trp Ser Tyr Pro Thr Ser Asn Asp Phe Gly Ser Glu Ile
            260                 265                 270

ata cca gca gct att gac gat tac aat gtc caa gca tac att ttc aaa   864
Ile Pro Ala Ala Ile Asp Asp Tyr Asn Val Gln Ala Tyr Ile Phe Lys
        275                 280                 285

gac tat tgg gaa gac att gga aca att aaa tcg ttt tat aat gct agc   912
Asp Tyr Trp Glu Asp Ile Gly Thr Ile Lys Ser Phe Tyr Asn Ala Ser
    290                 295                 300

ttg gca ctc aca caa gag ttt cca gag ttc caa ttt tac gat cca aaa   960
Leu Ala Leu Thr Gln Glu Phe Pro Glu Phe Gln Phe Tyr Asp Pro Lys
305                 310                 315                 320

aca cct ttt tac aca tct cct agg ttc ctt cca cca acc aag ata gac  1008
Thr Pro Phe Tyr Thr Ser Pro Arg Phe Leu Pro Pro Thr Lys Ile Asp
                325                 330                 335

aat tgc aag att aag gat gcc ata atc tct cat gga tgt ttc ttg cga  1056
Asn Cys Lys Ile Lys Asp Ala Ile Ile Ser His Gly Cys Phe Leu Arg
```

```
            340                 345                 350

gat tgt tct gtg gaa cac tcc ata gtg ggt gaa aga tcg cgc tta gat   1104
Asp Cys Ser Val Glu His Ser Ile Val Gly Glu Arg Ser Arg Leu Asp
        355                 360                 365

tgt ggt gtt gaa ctg aag gat act ttc atg atg gga gca gac tac tac   1152
Cys Gly Val Glu Leu Lys Asp Thr Phe Met Met Gly Ala Asp Tyr Tyr
    370                 375                 380

caa aca gaa tct gag att gcc tcc ctg tta gca gag ggg aaa gta ccg   1200
Gln Thr Glu Ser Glu Ile Ala Ser Leu Leu Ala Glu Gly Lys Val Pro
385                 390                 395                 400

att gga att ggg gaa aat aca aaa ata agg aaa tgt atc att gac aag   1248
Ile Gly Ile Gly Glu Asn Thr Lys Ile Arg Lys Cys Ile Ile Asp Lys
                405                 410                 415

aac gca aag ata gga aag aat gtt tca atc ata aat aaa gac ggt gtt   1296
Asn Ala Lys Ile Gly Lys Asn Val Ser Ile Ile Asn Lys Asp Gly Val
            420                 425                 430

caa gag gca gac cga cca gag gaa gga ttc tac ata cga tca ggg ata   1344
Gln Glu Ala Asp Arg Pro Glu Glu Gly Phe Tyr Ile Arg Ser Gly Ile
        435                 440                 445

atc att ata tta gag aaa gcc aca att aga gat gga aca gtc ata tga   1392
Ile Ile Ile Leu Glu Lys Ala Thr Ile Arg Asp Gly Thr Val Ile
    450                 455                 460
```

<210> 4
<211> 463
<212> PRT
<213> Solanum tuberosum

<400> 4

```
Met Ala Ser Val Ile Thr Thr Glu Asn Asp Thr Gln Thr Val Phe Val
  1               5                  10                  15

Asp Met Pro Arg Leu Glu Arg Arg Arg Ala Asn Pro Lys Asp Val Ala
             20                  25                  30

Ala Val Ile Leu Gly Gly Gly Glu Gly Thr Lys Leu Phe Pro Leu Thr
         35                  40                  45

Ser Arg Thr Ala Thr Pro Ala Val Pro Val Gly Gly Cys Tyr Arg Leu
     50                  55                  60
```

```
Ile Asp Ile Pro Met Ser Asn Cys Ile Asn Ser Ala Ile Asn Lys Ile
 65                  70                  75                  80

Phe Val Leu Thr Gln Tyr Asn Ser Ala Pro Leu Asn Arg His Ile Ala
                85                  90                  95

Arg Thr Tyr Phe Gly Asn Asn Val Ser Phe Gly Asp Gly Phe Val Glu
            100                 105                 110

Val Leu Ala Ala Thr Gln Thr Pro Gly Glu Ala Gly Lys Lys Trp Phe
        115                 120                 125

Gln Gly Thr Ala Asp Ala Val Arg Lys Phe Ile Trp Val Phe Glu Asp
    130                 135                 140

Ala Lys Asn Lys Asn Ile Glu Asn Ile Val Val Leu Ser Gly Asp His
145                 150                 155                 160

Leu Tyr Arg Met Asp Tyr Met Glu Leu Val Gln Asn His Ile Asp Arg
                165                 170                 175

Asn Ala Asp Ile Thr Leu Ser Cys Ala Pro Ala Glu Asp Ser Arg Ala
            180                 185                 190

Ser Asp Phe Gly Leu Val Lys Ile Asp Ser Arg Gly Arg Val Val Gln
        195                 200                 205

Phe Ala Glu Lys Pro Lys Gly Phe Asp Leu Lys Ala Met Gln Val Asp
    210                 215                 220

Thr Thr Leu Val Gly Leu Ser Pro Gln Asp Ala Lys Lys Ser Pro Tyr
225                 230                 235                 240

Ile Ala Ser Met Gly Val Tyr Val Phe Lys Thr Asp Val Leu Leu Lys
                245                 250                 255

Leu Leu Lys Trp Ser Tyr Pro Thr Ser Asn Asp Phe Gly Ser Glu Ile
            260                 265                 270

Ile Pro Ala Ala Ile Asp Asp Tyr Asn Val Gln Ala Tyr Ile Phe Lys
        275                 280                 285

Asp Tyr Trp Glu Asp Ile Gly Thr Ile Lys Ser Phe Tyr Asn Ala Ser
    290                 295                 300

Leu Ala Leu Thr Gln Glu Phe Pro Glu Phe Gln Phe Tyr Asp Pro Lys
305                 310                 315                 320
```

```
Thr Pro Phe Tyr Thr Ser Pro Arg Phe Leu Pro Pro Thr Lys Ile Asp
                325                 330                 335

Asn Cys Lys Ile Lys Asp Ala Ile Ile Ser His Gly Cys Phe Leu Arg
            340                 345                 350

Asp Cys Ser Val Glu His Ser Ile Val Gly Glu Arg Ser Arg Leu Asp
        355                 360                 365

Cys Gly Val Glu Leu Lys Asp Thr Phe Met Met Gly Ala Asp Tyr Tyr
    370                 375                 380

Gln Thr Glu Ser Glu Ile Ala Ser Leu Leu Ala Glu Gly Lys Val Pro
385                 390                 395                 400

Ile Gly Ile Gly Glu Asn Thr Lys Ile Arg Lys Cys Ile Ile Asp Lys
                405                 410                 415

Asn Ala Lys Ile Gly Lys Asn Val Ser Ile Ile Asn Lys Asp Gly Val
            420                 425                 430

Gln Glu Ala Asp Arg Pro Glu Glu Gly Phe Tyr Ile Arg Ser Gly Ile
        435                 440                 445

Ile Ile Ile Leu Glu Lys Ala Thr Ile Arg Asp Gly Thr Val Ile
    450                 455                 460
```

<210> 5
<211> 1341
<212> DNA
<213> Solanum tuberosum

<220>
<221> CDS
<222> (1)..(1341)

<400> 5

```
atg gcc cgt ctt gag aga cgc cgg gca aat cca aag gat gtg gct gca   48
Met Ala Arg Leu Glu Arg Arg Arg Ala Asn Pro Lys Asp Val Ala Ala
  1               5                  10                  15

gtc ata ctg gga gga gga gaa ggg acc aag tta ttc cca ctt aca agt   96
Val Ile Leu Gly Gly Gly Glu Gly Thr Lys Leu Phe Pro Leu Thr Ser
            20                  25                  30

aga act gca acc cct gct gtt ccg gtt gga gga tgc tac agg cta ata   144
```

```
Arg Thr Ala Thr Pro Ala Val Pro Val Gly Gly Cys Tyr Arg Leu Ile
         35                  40                  45

gac atc cca atg agc aac tgt atc aac agt gct att aac aag att ttt   192
Asp Ile Pro Met Ser Asn Cys Ile Asn Ser Ala Ile Asn Lys Ile Phe
     50                  55                  60

gtg ctg aca cag tac aat tct gct ccc ctg aat cgt cac att gct cga   240
Val Leu Thr Gln Tyr Asn Ser Ala Pro Leu Asn Arg His Ile Ala Arg
 65                  70                  75                  80

aca tat ttt ggc aat ggt gtg agc ttt gga gat gga ttt gtc gag gta   288
Thr Tyr Phe Gly Asn Gly Val Ser Phe Gly Asp Gly Phe Val Glu Val
                 85                  90                  95

cta gct gca act cag aca ccc ggg gaa gca gga aaa aaa tgg ttt caa   336
Leu Ala Ala Thr Gln Thr Pro Gly Glu Ala Gly Lys Lys Trp Phe Gln
            100                 105                 110

gga aca gca gat gct gtt aga aaa ttt ata tgg gtt ttt gag gac gct   384
Gly Thr Ala Asp Ala Val Arg Lys Phe Ile Trp Val Phe Glu Asp Ala
        115                 120                 125

aag aac aag aat att gaa aat atc gtt gta cta tct ggg gat cat ctt   432
Lys Asn Lys Asn Ile Glu Asn Ile Val Val Leu Ser Gly Asp His Leu
    130                 135                 140

tat agg atg gat tat atg gag ttg gtg cag aac cat att gac agg aat   480
Tyr Arg Met Asp Tyr Met Glu Leu Val Gln Asn His Ile Asp Arg Asn
145                 150                 155                 160

gct gat att act ctt tca tgt gca cca gct gag gac agc cga gca tca   528
Ala Asp Ile Thr Leu Ser Cys Ala Pro Ala Glu Asp Ser Arg Ala Ser
                165                 170                 175

gat ttt ggg ctg gtc aag att gac agc aga ggc aga gta gtc cag ttt   576
Asp Phe Gly Leu Val Lys Ile Asp Ser Arg Gly Arg Val Val Gln Phe
            180                 185                 190

gct gaa aaa cca aaa ggt ttt gat ctt aaa gca atg caa gta gat act   624
Ala Glu Lys Pro Lys Gly Phe Asp Leu Lys Ala Met Gln Val Asp Thr
        195                 200                 205

act ctt gtt gga tta tct cca caa gat gcg aag aaa tcc ccc tat att   672
Thr Leu Val Gly Leu Ser Pro Gln Asp Ala Lys Lys Ser Pro Tyr Ile
    210                 215                 220

gct tca atg gga gtt tat gta ttc aag aca gat gta ttg ttg aag ctc   720
```

```
Ala Ser Met Gly Val Tyr Val Phe Lys Thr Asp Val Leu Leu Lys Leu
225                 230                 235                 240

ttg aaa tgg agc tat ccc act tct aat gat ttt ggc tct gaa att ata   768
Leu Lys Trp Ser Tyr Pro Thr Ser Asn Asp Phe Gly Ser Glu Ile Ile
                245                 250                 255

cca gca gct att gac gat tac aat gtc caa gca tac att ttc aaa gac   816
Pro Ala Ala Ile Asp Asp Tyr Asn Val Gln Ala Tyr Ile Phe Lys Asp
            260                 265                 270

tat tgg gaa gac att gga aca att aaa tcg ttt tat aat gct agc ttg   864
Tyr Trp Glu Asp Ile Gly Thr Ile Lys Ser Phe Tyr Asn Ala Ser Leu
        275                 280                 285

gca ctc aca caa gag ttt cca gag ttc caa ttt tac gat cca aaa aca   912
Ala Leu Thr Gln Glu Phe Pro Glu Phe Gln Phe Tyr Asp Pro Lys Thr
    290                 295                 300

cct ttt tac aca tct cct agg ttc ctt cca cca acc aag ata gac aat   960
Pro Phe Tyr Thr Ser Pro Arg Phe Leu Pro Pro Thr Lys Ile Asp Asn
305                 310                 315                 320

tgc aag att aag gat gcc ata atc tct cat gga tgt ttc ttg cga gat   1008
Cys Lys Ile Lys Asp Ala Ile Ile Ser His Gly Cys Phe Leu Arg Asp
                325                 330                 335

tgt tct gtg gaa cac tcc ata gtg ggt gaa aga tcg cgc tta gat tgt   1056
Cys Ser Val Glu His Ser Ile Val Gly Glu Arg Ser Arg Leu Asp Cys
            340                 345                 350

ggt gtt gaa ctg aag gat act ttc atg atg gga gca gac tac tac caa   1104
Gly Val Glu Leu Lys Asp Thr Phe Met Met Gly Ala Asp Tyr Tyr Gln
        355                 360                 365

aca gaa tct gag att gcc tcc ctg tta gca gag ggg aaa gta ccg att   1152
Thr Glu Ser Glu Ile Ala Ser Leu Leu Ala Glu Gly Lys Val Pro Ile
    370                 375                 380

gga att ggg gaa aat aca aaa ata agg aaa tgt atc att gac aag aac   1200
Gly Ile Gly Glu Asn Thr Lys Ile Arg Lys Cys Ile Ile Asp Lys Asn
385                 390                 395                 400

gca aag ata gga aag aat gtt tca atc ata aat aaa gac ggt gtt caa   1248
Ala Lys Ile Gly Lys Asn Val Ser Ile Ile Asn Lys Asp Gly Val Gln
                405                 410                 415

gag gca gac cga cca gag gaa gga ttc tac ata cga tca ggg ata atc   1296
```

```
Glu Ala Asp Arg Pro Glu Glu Gly Phe Tyr Ile Arg Ser Gly Ile Ile
            420                 425                 430

att ata tta gag aaa gcc aca att aga gat gga aca gtc ata tga       1341
Ile Ile Leu Glu Lys Ala Thr Ile Arg Asp Gly Thr Val Ile
        435                 440                 445
```

<210> 6
<211> 446
<212> PRT
<213> Solanum tuberosum

<400> 6

```
Met Ala Arg Leu Glu Arg Arg Arg Ala Asn Pro Lys Asp Val Ala Ala
  1               5                  10                  15

Val Ile Leu Gly Gly Gly Glu Gly Thr Lys Leu Phe Pro Leu Thr Ser
            20                  25                  30

Arg Thr Ala Thr Pro Ala Val Pro Val Gly Gly Cys Tyr Arg Leu Ile
        35                  40                  45

Asp Ile Pro Met Ser Asn Cys Ile Asn Ser Ala Ile Asn Lys Ile Phe
    50                  55                  60

Val Leu Thr Gln Tyr Asn Ser Ala Pro Leu Asn Arg His Ile Ala Arg
65                  70                  75                  80

Thr Tyr Phe Gly Asn Gly Val Ser Phe Gly Asp Gly Phe Val Glu Val
                85                  90                  95

Leu Ala Ala Thr Gln Thr Pro Gly Glu Ala Gly Lys Lys Trp Phe Gln
            100                 105                 110

Gly Thr Ala Asp Ala Val Arg Lys Phe Ile Trp Val Phe Glu Asp Ala
        115                 120                 125

Lys Asn Lys Asn Ile Glu Asn Ile Val Val Leu Ser Gly Asp His Leu
    130                 135                 140

Tyr Arg Met Asp Tyr Met Glu Leu Val Gln Asn His Ile Asp Arg Asn
145                 150                 155                 160

Ala Asp Ile Thr Leu Ser Cys Ala Pro Ala Glu Asp Ser Arg Ala Ser
                165                 170                 175

Asp Phe Gly Leu Val Lys Ile Asp Ser Arg Gly Arg Val Val Gln Phe
```

```
            180                 185                 190

Ala Glu Lys Pro Lys Gly Phe Asp Leu Lys Ala Met Gln Val Asp Thr
        195                 200                 205

Thr Leu Val Gly Leu Ser Pro Gln Asp Ala Lys Lys Ser Pro Tyr Ile
    210                 215                 220

Ala Ser Met Gly Val Tyr Val Phe Lys Thr Asp Val Leu Leu Lys Leu
225                 230                 235                 240

Leu Lys Trp Ser Tyr Pro Thr Ser Asn Asp Phe Gly Ser Glu Ile Ile
                245                 250                 255

Pro Ala Ala Ile Asp Asp Tyr Asn Val Gln Ala Tyr Ile Phe Lys Asp
            260                 265                 270

Tyr Trp Glu Asp Ile Gly Thr Ile Lys Ser Phe Tyr Asn Ala Ser Leu
        275                 280                 285

Ala Leu Thr Gln Glu Phe Pro Glu Phe Gln Phe Tyr Asp Pro Lys Thr
    290                 295                 300

Pro Phe Tyr Thr Ser Pro Arg Phe Leu Pro Pro Thr Lys Ile Asp Asn
305                 310                 315                 320

Cys Lys Ile Lys Asp Ala Ile Ile Ser His Gly Cys Phe Leu Arg Asp
                325                 330                 335

Cys Ser Val Glu His Ser Ile Val Gly Glu Arg Ser Arg Leu Asp Cys
            340                 345                 350

Gly Val Glu Leu Lys Asp Thr Phe Met Met Gly Ala Asp Tyr Tyr Gln
        355                 360                 365

Thr Glu Ser Glu Ile Ala Ser Leu Leu Ala Glu Gly Lys Val Pro Ile
    370                 375                 380

Gly Ile Gly Glu Asn Thr Lys Ile Arg Lys Cys Ile Ile Asp Lys Asn
385                 390                 395                 400

Ala Lys Ile Gly Lys Asn Val Ser Ile Ile Asn Lys Asp Gly Val Gln
                405                 410                 415

Glu Ala Asp Arg Pro Glu Glu Gly Phe Tyr Ile Arg Ser Gly Ile Ile
            420                 425                 430

Ile Ile Leu Glu Lys Ala Thr Ile Arg Asp Gly Thr Val Ile
```

435 440 445

<210> 7
<211> 1392
<212> DNA
<213> Solanum tuberosum

<220>
<221> CDS
<222> (1)..(1392)

<400> 7
atg gcc tct gtg atc act act gaa aat gac aca cag act gtg ttc gta 48
Met Ala Ser Val Ile Thr Thr Glu Asn Asp Thr Gln Thr Val Phe Val
1 5 10 15

gat atg cca cgt ctt gag aga cgc cgg gca aat cca aag gat gtg gct 96
Asp Met Pro Arg Leu Glu Arg Arg Arg Ala Asn Pro Lys Asp Val Ala
20 25 30

gca gtc ata ctg gga gga gga aaa ggg acc aag tta ttc cca ctt aca 144
Ala Val Ile Leu Gly Gly Gly Lys Gly Thr Lys Leu Phe Pro Leu Thr
35 40 45

agt aga act gca acc ctt gct gtt ccg gtt gga gga tgc tac agg cta 192
Ser Arg Thr Ala Thr Leu Ala Val Pro Val Gly Gly Cys Tyr Arg Leu
50 55 60

ata gac atc cca atg agc aac tgt atc aac agt gct att aac aag att 240
Ile Asp Ile Pro Met Ser Asn Cys Ile Asn Ser Ala Ile Asn Lys Ile
65 70 75 80

ttt gtg ctg aca cag tac aat tct gct ccc ctg aat cgt cac att gct 288
Phe Val Leu Thr Gln Tyr Asn Ser Ala Pro Leu Asn Arg His Ile Ala
85 90 95

cga aca tat ttt ggc aat ggt gtg agc ttt gga gat gga ttt gtc gag 336
Arg Thr Tyr Phe Gly Asn Gly Val Ser Phe Gly Asp Gly Phe Val Glu
100 105 110

gta cta gct gca act cag aca ccc ggg gaa gca gga aaa aaa tgg ttt 384
Val Leu Ala Ala Thr Gln Thr Pro Gly Glu Ala Gly Lys Lys Trp Phe
115 120 125

caa gga aca gca gat gct gtt aga aaa ttt ata tgg gtt ttt gag gac 432
Gln Gly Thr Ala Asp Ala Val Arg Lys Phe Ile Trp Val Phe Glu Asp
130 135 140

```
gct aag aac aag aat att gaa aat atc gtt gta cta tct ggg gat cat   480
Ala Lys Asn Lys Asn Ile Glu Asn Ile Val Val Leu Ser Gly Asp His
145                 150                 155                 160

ctt tat agg atg gat tat atg gag ttg gtg cag aac cat att gac agg   528
Leu Tyr Arg Met Asp Tyr Met Glu Leu Val Gln Asn His Ile Asp Arg
                165                 170                 175

aat gct gat att act ctt tca tgt gca cca gct gag gac agc cga gca   576
Asn Ala Asp Ile Thr Leu Ser Cys Ala Pro Ala Glu Asp Ser Arg Ala
            180                 185                 190

tca gat ttt ggg ctg gtc aag att gac agc aga ggc aga gta gtc cag   624
Ser Asp Phe Gly Leu Val Lys Ile Asp Ser Arg Gly Arg Val Val Gln
        195                 200                 205

ttt gct gaa aaa cca aaa ggt ttt gat ctt aaa gca atg caa gta gat   672
Phe Ala Glu Lys Pro Lys Gly Phe Asp Leu Lys Ala Met Gln Val Asp
    210                 215                 220

act act ctt gtt gga tta tct cca caa gat gcg aag aaa tcc ccc tat   720
Thr Thr Leu Val Gly Leu Ser Pro Gln Asp Ala Lys Lys Ser Pro Tyr
225                 230                 235                 240

att gct tca atg gga gtt tat gta ttc aag aca gat gta ttg ttg aag   768
Ile Ala Ser Met Gly Val Tyr Val Phe Lys Thr Asp Val Leu Leu Lys
                245                 250                 255

ctc ttg aaa tgg agc tat ccc act tct aat gat ttt ggc tct gaa att   816
Leu Leu Lys Trp Ser Tyr Pro Thr Ser Asn Asp Phe Gly Ser Glu Ile
            260                 265                 270

ata cca gca gct att gac gat tac aat gtc caa gca tac att ttc aaa   864
Ile Pro Ala Ala Ile Asp Asp Tyr Asn Val Gln Ala Tyr Ile Phe Lys
        275                 280                 285

gac tat tgg gaa gac att gga aca att aaa tcg ttt tat aat gct agc   912
Asp Tyr Trp Glu Asp Ile Gly Thr Ile Lys Ser Phe Tyr Asn Ala Ser
    290                 295                 300

ttg gca ctc aca caa gag ttt cca gag ttc caa ttt tac gat cca aaa   960
Leu Ala Leu Thr Gln Glu Phe Pro Glu Phe Gln Phe Tyr Asp Pro Lys
305                 310                 315                 320

aca cct ttt tac aca tct cct agg ttc ctt cca cca acc aag ata gac  1008
Thr Pro Phe Tyr Thr Ser Pro Arg Phe Leu Pro Pro Thr Lys Ile Asp
                325                 330                 335
```

```
aat tgc aag att aag gat gcc ata atc tct cat gga tgt ttc ttg cga   1056
Asn Cys Lys Ile Lys Asp Ala Ile Ile Ser His Gly Cys Phe Leu Arg
            340                 345                 350

gat tgt tct gtg gaa cac tcc ata gtg ggt gaa aga tcg cgc tta gat   1104
Asp Cys Ser Val Glu His Ser Ile Val Gly Glu Arg Ser Arg Leu Asp
        355                 360                 365

tgt ggt gtt gaa ctg aag gat act ttc atg atg gga gca gac tac tac   1152
Cys Gly Val Glu Leu Lys Asp Thr Phe Met Met Gly Ala Asp Tyr Tyr
    370                 375                 380

caa aca gaa tct gag att gcc tcc ctg tta gca gag ggg aaa gta ccg   1200
Gln Thr Glu Ser Glu Ile Ala Ser Leu Leu Ala Glu Gly Lys Val Pro
385                 390                 395                 400

att gga att ggg gaa aat aca aaa ata agg aaa tgt atc att gac aag   1248
Ile Gly Ile Gly Glu Asn Thr Lys Ile Arg Lys Cys Ile Ile Asp Lys
                405                 410                 415

aac gca aag ata gga aag aat gtt tca atc ata aat aaa gac ggt gtt   1296
Asn Ala Lys Ile Gly Lys Asn Val Ser Ile Ile Asn Lys Asp Gly Val
            420                 425                 430

caa gag gca gac cga cca gag gaa gga ttc tac ata cga tca ggg ata   1344
Gln Glu Ala Asp Arg Pro Glu Glu Gly Phe Tyr Ile Arg Ser Gly Ile
        435                 440                 445

atc att ata tta gag aaa gcc aca att aga gat gga aca gtc ata tga   1392
Ile Ile Ile Leu Glu Lys Ala Thr Ile Arg Asp Gly Thr Val Ile
    450                 455                 460


<210> 8
<211> 463
<212> PRT
<213> Solanum tuberosum

<400> 8
Met Ala Ser Val Ile Thr Thr Glu Asn Asp Thr Gln Thr Val Phe Val
  1               5                  10                  15

Asp Met Pro Arg Leu Glu Arg Arg Arg Ala Asn Pro Lys Asp Val Ala
             20                  25                  30

Ala Val Ile Leu Gly Gly Gly Lys Gly Thr Lys Leu Phe Pro Leu Thr
         35                  40                  45
```

```
Ser Arg Thr Ala Thr Leu Ala Val Pro Val Gly Gly Cys Tyr Arg Leu
     50                  55                  60

Ile Asp Ile Pro Met Ser Asn Cys Ile Asn Ser Ala Ile Asn Lys Ile
 65                  70                  75                  80

Phe Val Leu Thr Gln Tyr Asn Ser Ala Pro Leu Asn Arg His Ile Ala
                85                  90                  95

Arg Thr Tyr Phe Gly Asn Gly Val Ser Phe Gly Asp Gly Phe Val Glu
            100                 105                 110

Val Leu Ala Ala Thr Gln Thr Pro Gly Glu Ala Gly Lys Lys Trp Phe
        115                 120                 125

Gln Gly Thr Ala Asp Ala Val Arg Lys Phe Ile Trp Val Phe Glu Asp
    130                 135                 140

Ala Lys Asn Lys Asn Ile Glu Asn Ile Val Val Leu Ser Gly Asp His
145                 150                 155                 160

Leu Tyr Arg Met Asp Tyr Met Glu Leu Val Gln Asn His Ile Asp Arg
                165                 170                 175

Asn Ala Asp Ile Thr Leu Ser Cys Ala Pro Ala Glu Asp Ser Arg Ala
            180                 185                 190

Ser Asp Phe Gly Leu Val Lys Ile Asp Ser Arg Gly Arg Val Val Gln
        195                 200                 205

Phe Ala Glu Lys Pro Lys Gly Phe Asp Leu Lys Ala Met Gln Val Asp
    210                 215                 220

Thr Thr Leu Val Gly Leu Ser Pro Gln Asp Ala Lys Lys Ser Pro Tyr
225                 230                 235                 240

Ile Ala Ser Met Gly Val Tyr Val Phe Lys Thr Asp Val Leu Leu Lys
                245                 250                 255

Leu Leu Lys Trp Ser Tyr Pro Thr Ser Asn Asp Phe Gly Ser Glu Ile
            260                 265                 270

Ile Pro Ala Ala Ile Asp Asp Tyr Asn Val Gln Ala Tyr Ile Phe Lys
        275                 280                 285

Asp Tyr Trp Glu Asp Ile Gly Thr Ile Lys Ser Phe Tyr Asn Ala Ser
    290                 295                 300
```

```
Leu Ala Leu Thr Gln Glu Phe Pro Glu Phe Gln Phe Tyr Asp Pro Lys
305                 310                 315                 320

Thr Pro Phe Tyr Thr Ser Pro Arg Phe Leu Pro Pro Thr Lys Ile Asp
                325                 330                 335

Asn Cys Lys Ile Lys Asp Ala Ile Ile Ser His Gly Cys Phe Leu Arg
            340                 345                 350

Asp Cys Ser Val Glu His Ser Ile Val Gly Glu Arg Ser Arg Leu Asp
        355                 360                 365

Cys Gly Val Glu Leu Lys Asp Thr Phe Met Met Gly Ala Asp Tyr Tyr
    370                 375                 380

Gln Thr Glu Ser Glu Ile Ala Ser Leu Leu Ala Glu Gly Lys Val Pro
385                 390                 395                 400

Ile Gly Ile Gly Glu Asn Thr Lys Ile Arg Lys Cys Ile Ile Asp Lys
                405                 410                 415

Asn Ala Lys Ile Gly Lys Asn Val Ser Ile Ile Asn Lys Asp Gly Val
            420                 425                 430

Gln Glu Ala Asp Arg Pro Glu Glu Gly Phe Tyr Ile Arg Ser Gly Ile
        435                 440                 445

Ile Ile Ile Leu Glu Lys Ala Thr Ile Arg Asp Gly Thr Val Ile
    450                 455                 460
```

<210> 9
<211> 1392
<212> DNA
<213> Solanum tuberosum

<220>
<221> CDS
<222> (1)..(1392)

<400> 9

```
atg gcc tct gtg atc act act gaa aat gac aca cag act gtg ttc gta   48
Met Ala Ser Val Ile Thr Thr Glu Asn Asp Thr Gln Thr Val Phe Val
  1               5                  10                  15

gat atg cca cgt ctt gag aga cgc cgg gca aat cca aag gat gtg gct   96
Asp Met Pro Arg Leu Glu Arg Arg Arg Ala Asn Pro Lys Asp Val Ala
```

```
            20                  25                  30

gca gtc ata ctg gga gga gga gaa ggg acc aag tta ttc cca ctt aca   144
Ala Val Ile Leu Gly Gly Gly Glu Gly Thr Lys Leu Phe Pro Leu Thr
        35                  40                  45

agt aga act gca acc ctt gct gtt ccg gtt gga gga tgc tac agg cta   192
Ser Arg Thr Ala Thr Leu Ala Val Pro Val Gly Gly Cys Tyr Arg Leu
    50                  55                  60

ata gac atc cca atg agc aac tgt atc aac agt gct att aac aag att   240
Ile Asp Ile Pro Met Ser Asn Cys Ile Asn Ser Ala Ile Asn Lys Ile
65                  70                  75                  80

ttt gtg ctg aca cag tac aat tct gct ccc ctg aat cgt cac att gct   288
Phe Val Leu Thr Gln Tyr Asn Ser Ala Pro Leu Asn Arg His Ile Ala
                85                  90                  95

cga aca tat ttt ggc aat aat gtg agc ttt gga gat gga ttt gtc gag   336
Arg Thr Tyr Phe Gly Asn Asn Val Ser Phe Gly Asp Gly Phe Val Glu
            100                 105                 110

gta cta gct gca act cag aca ccc ggg gaa gca gga aaa aaa tgg ttt   384
Val Leu Ala Ala Thr Gln Thr Pro Gly Glu Ala Gly Lys Lys Trp Phe
        115                 120                 125

caa gga aca gca gat gct gtt aga aaa ttt ata tgg gtt ttt gag gac   432
Gln Gly Thr Ala Asp Ala Val Arg Lys Phe Ile Trp Val Phe Glu Asp
    130                 135                 140

gct aag aac aag aat att gaa aat atc gtt gta cta tct ggg gat cat   480
Ala Lys Asn Lys Asn Ile Glu Asn Ile Val Val Leu Ser Gly Asp His
145                 150                 155                 160

ctt tat agg atg gat tat atg gag ttg gtg cag aac cat att gac agg   528
Leu Tyr Arg Met Asp Tyr Met Glu Leu Val Gln Asn His Ile Asp Arg
                165                 170                 175

aat gct gat att act ctt tca tgt gca cca gct gag gac agc cga gca   576
Asn Ala Asp Ile Thr Leu Ser Cys Ala Pro Ala Glu Asp Ser Arg Ala
            180                 185                 190

tca gat ttt ggg ctg gtc aag att gac agc aga ggc aga gta gtc cag   624
Ser Asp Phe Gly Leu Val Lys Ile Asp Ser Arg Gly Arg Val Val Gln
        195                 200                 205

ttt gct gaa aaa cca aaa ggt ttt gat ctt aaa gca atg caa gta gat   672
Phe Ala Glu Lys Pro Lys Gly Phe Asp Leu Lys Ala Met Gln Val Asp
```

```
    210                 215                 220

act act ctt gtt gga tta tct cca caa gat gcg aag aaa tcc ccc tat   720
Thr Thr Leu Val Gly Leu Ser Pro Gln Asp Ala Lys Lys Ser Pro Tyr
225                 230                 235                 240

att gct tca atg gga gtt tat gta ttc aag aca gat gta ttg ttg aag   768
Ile Ala Ser Met Gly Val Tyr Val Phe Lys Thr Asp Val Leu Leu Lys
                245                 250                 255

ctc ttg aaa tgg agc tat ccc act tct aat gat ttt ggc tct gaa att   816
Leu Leu Lys Trp Ser Tyr Pro Thr Ser Asn Asp Phe Gly Ser Glu Ile
            260                 265                 270

ata cca gca gct att gac gat tac aat gtc caa gca tac att ttc aaa   864
Ile Pro Ala Ala Ile Asp Asp Tyr Asn Val Gln Ala Tyr Ile Phe Lys
        275                 280                 285

gac tat tgg gaa gac att gga aca att aaa tcg ttt tat aat gct agc   912
Asp Tyr Trp Glu Asp Ile Gly Thr Ile Lys Ser Phe Tyr Asn Ala Ser
    290                 295                 300

ttg gca ctc aca caa gag ttt cca gag ttc caa ttt tac gat cca aaa   960
Leu Ala Leu Thr Gln Glu Phe Pro Glu Phe Gln Phe Tyr Asp Pro Lys
305                 310                 315                 320

aca cct ttt tac aca tct cct agg ttc ctt cca cca acc aag ata gac   1008
Thr Pro Phe Tyr Thr Ser Pro Arg Phe Leu Pro Pro Thr Lys Ile Asp
                325                 330                 335

aat tgc aag att aag gat gcc ata atc tct cat gga tgt ttc ttg cga   1056
Asn Cys Lys Ile Lys Asp Ala Ile Ile Ser His Gly Cys Phe Leu Arg
            340                 345                 350

gat tgt tct gtg gaa cac tcc ata gtg ggt gaa aga tcg cgc tta gat   1104
Asp Cys Ser Val Glu His Ser Ile Val Gly Glu Arg Ser Arg Leu Asp
        355                 360                 365

tgt ggt gtt gaa ctg aag gat act ttc atg atg gga gca gac tac tac   1152
Cys Gly Val Glu Leu Lys Asp Thr Phe Met Met Gly Ala Asp Tyr Tyr
    370                 375                 380

caa aca gaa tct gag att gcc tcc ctg tta gca gag ggg aaa gta ccg   1200
Gln Thr Glu Ser Glu Ile Ala Ser Leu Leu Ala Glu Gly Lys Val Pro
385                 390                 395                 400

att gga att ggg gaa aat aca aaa ata agg aaa tgt atc att gac aag   1248
Ile Gly Ile Gly Glu Asn Thr Lys Ile Arg Lys Cys Ile Ile Asp Lys
```

```
                405                 410                 415

aac gca aag ata gga aag aat gtt tca atc ata aat aaa gac ggt gtt   1296
Asn Ala Lys Ile Gly Lys Asn Val Ser Ile Ile Asn Lys Asp Gly Val
            420                 425                 430

caa gag gca gac cga cca gag gaa gga ttc tac ata cga tca ggg ata   1344
Gln Glu Ala Asp Arg Pro Glu Glu Gly Phe Tyr Ile Arg Ser Gly Ile
        435                 440                 445

atc att ata tta gag aaa gcc aca att aga gat gga aca gtc ata tga   1392
Ile Ile Ile Leu Glu Lys Ala Thr Ile Arg Asp Gly Thr Val Ile
    450                 455                 460


<210> 10
<211> 463
<212> PRT
<213> Solanum tuberosum

<400> 10
Met Ala Ser Val Ile Thr Thr Glu Asn Asp Thr Gln Thr Val Phe Val
  1               5                  10                  15

Asp Met Pro Arg Leu Glu Arg Arg Arg Ala Asn Pro Lys Asp Val Ala
            20                  25                  30

Ala Val Ile Leu Gly Gly Gly Glu Gly Thr Lys Leu Phe Pro Leu Thr
        35                  40                  45

Ser Arg Thr Ala Thr Leu Ala Val Pro Val Gly Gly Cys Tyr Arg Leu
    50                  55                  60

Ile Asp Ile Pro Met Ser Asn Cys Ile Asn Ser Ala Ile Asn Lys Ile
65                  70                  75                  80

Phe Val Leu Thr Gln Tyr Asn Ser Ala Pro Leu Asn Arg His Ile Ala
                85                  90                  95

Arg Thr Tyr Phe Gly Asn Asn Val Ser Phe Gly Asp Gly Phe Val Glu
            100                 105                 110

Val Leu Ala Ala Thr Gln Thr Pro Gly Glu Ala Gly Lys Lys Trp Phe
        115                 120                 125

Gln Gly Thr Ala Asp Ala Val Arg Lys Phe Ile Trp Val Phe Glu Asp
    130                 135                 140
```

```
Ala Lys Asn Lys Asn Ile Glu Asn Ile Val Val Leu Ser Gly Asp His
145                 150                 155                 160

Leu Tyr Arg Met Asp Tyr Met Glu Leu Val Gln Asn His Ile Asp Arg
                165                 170                 175

Asn Ala Asp Ile Thr Leu Ser Cys Ala Pro Ala Glu Asp Ser Arg Ala
            180                 185                 190

Ser Asp Phe Gly Leu Val Lys Ile Asp Ser Arg Gly Arg Val Val Gln
        195                 200                 205

Phe Ala Glu Lys Pro Lys Gly Phe Asp Leu Lys Ala Met Gln Val Asp
    210                 215                 220

Thr Thr Leu Val Gly Leu Ser Pro Gln Asp Ala Lys Lys Ser Pro Tyr
225                 230                 235                 240

Ile Ala Ser Met Gly Val Tyr Val Phe Lys Thr Asp Val Leu Leu Lys
                245                 250                 255

Leu Leu Lys Trp Ser Tyr Pro Thr Ser Asn Asp Phe Gly Ser Glu Ile
            260                 265                 270

Ile Pro Ala Ala Ile Asp Asp Tyr Asn Val Gln Ala Tyr Ile Phe Lys
        275                 280                 285

Asp Tyr Trp Glu Asp Ile Gly Thr Ile Lys Ser Phe Tyr Asn Ala Ser
    290                 295                 300

Leu Ala Leu Thr Gln Glu Phe Pro Glu Phe Gln Phe Tyr Asp Pro Lys
305                 310                 315                 320

Thr Pro Phe Tyr Thr Ser Pro Arg Phe Leu Pro Pro Thr Lys Ile Asp
                325                 330                 335

Asn Cys Lys Ile Lys Asp Ala Ile Ile Ser His Gly Cys Phe Leu Arg
            340                 345                 350

Asp Cys Ser Val Glu His Ser Ile Val Gly Glu Arg Ser Arg Leu Asp
        355                 360                 365

Cys Gly Val Glu Leu Lys Asp Thr Phe Met Met Gly Ala Asp Tyr Tyr
    370                 375                 380

Gln Thr Glu Ser Glu Ile Ala Ser Leu Leu Ala Glu Gly Lys Val Pro
385                 390                 395                 400
```

```
Ile Gly Ile Gly Glu Asn Thr Lys Ile Arg Lys Cys Ile Ile Asp Lys
                405                 410                 415

Asn Ala Lys Ile Gly Lys Asn Val Ser Ile Ile Asn Lys Asp Gly Val
            420                 425                 430

Gln Glu Ala Asp Arg Pro Glu Glu Gly Phe Tyr Ile Arg Ser Gly Ile
        435                 440                 445

Ile Ile Ile Leu Glu Lys Ala Thr Ile Arg Asp Gly Thr Val Ile
    450                 455                 460


<210> 11
<211> 1392
<212> DNA
<213> Solanum tuberosum

<220>
<221> CDS
<222> (1)..(1392)

<400> 11
atg gcc tct gtg atc act act gaa aat gac aca cag act gtg ttc gta   48
Met Ala Ser Val Ile Thr Thr Glu Asn Asp Thr Gln Thr Val Phe Val
  1               5                  10                  15

gat atg cca cgt ctt gag aga cgc cgg gca aat cca aag gat gtg gct   96
Asp Met Pro Arg Leu Glu Arg Arg Arg Ala Asn Pro Lys Asp Val Ala
             20                  25                  30

gca gtc ata ctg gga gga gga gaa ggg acc aag tta ttc cca ctt aca   144
Ala Val Ile Leu Gly Gly Gly Glu Gly Thr Lys Leu Phe Pro Leu Thr
         35                  40                  45

agt aga act gca acc ctt gct gtt ccg gtt gga gga tgc tac agg cta   192
Ser Arg Thr Ala Thr Leu Ala Val Pro Val Gly Gly Cys Tyr Arg Leu
     50                  55                  60

ata gac atc cta atg agc aac tgt atc aac agt gct att aac aag att   240
Ile Asp Ile Leu Met Ser Asn Cys Ile Asn Ser Ala Ile Asn Lys Ile
 65                  70                  75                  80

ttt gtg ctg aca cag tac aat tct gct ccc ctg aat cgt cac att gct   288
Phe Val Leu Thr Gln Tyr Asn Ser Ala Pro Leu Asn Arg His Ile Ala
                 85                  90                  95

cga aca tat ttt ggc aat ggt gtg agc ttt gga gat gga ttt gtc gag   336
```

```
Arg Thr Tyr Phe Gly Asn Gly Val Ser Phe Gly Asp Gly Phe Val Glu
            100                 105                 110

gta cta gct gca act cag aca ccc ggg gaa gca gga aaa aaa tgg ttt   384
Val Leu Ala Ala Thr Gln Thr Pro Gly Glu Ala Gly Lys Lys Trp Phe
        115                 120                 125

caa gga aca gca gat gct gtt aga aaa ttt ata tgg gtt ttt gag gac   432
Gln Gly Thr Ala Asp Ala Val Arg Lys Phe Ile Trp Val Phe Glu Asp
    130                 135                 140

gct aag aac aag aat att gaa aat atc gtt gta cta tct ggg gat cat   480
Ala Lys Asn Lys Asn Ile Glu Asn Ile Val Val Leu Ser Gly Asp His
145                 150                 155                 160

ctt tat agg atg gat tat atg gag ttg gtg cag aac cat att gac agg   528
Leu Tyr Arg Met Asp Tyr Met Glu Leu Val Gln Asn His Ile Asp Arg
                165                 170                 175

aat gct gat att act ctt tca tgt gca cca gct gag gac agc cga gca   576
Asn Ala Asp Ile Thr Leu Ser Cys Ala Pro Ala Glu Asp Ser Arg Ala
            180                 185                 190

tca gat ttt ggg ctg gtc aag att gac agc aga ggc aga gta gtc cag   624
Ser Asp Phe Gly Leu Val Lys Ile Asp Ser Arg Gly Arg Val Val Gln
        195                 200                 205

ttt gct gaa aaa cca aaa ggt ttt gat ctt aaa gca atg caa gta gat   672
Phe Ala Glu Lys Pro Lys Gly Phe Asp Leu Lys Ala Met Gln Val Asp
    210                 215                 220

act act ctt gtt gga tta tct cca caa gat gcg aag aaa tcc ccc tat   720
Thr Thr Leu Val Gly Leu Ser Pro Gln Asp Ala Lys Lys Ser Pro Tyr
225                 230                 235                 240

att gct tca atg gga gtt tat gta ttc aag aca gat gta ttg ttg aag   768
Ile Ala Ser Met Gly Val Tyr Val Phe Lys Thr Asp Val Leu Leu Lys
                245                 250                 255

ctc ttg aaa tgg agc tat ccc act tct aat gat ttt ggc tct gaa att   816
Leu Leu Lys Trp Ser Tyr Pro Thr Ser Asn Asp Phe Gly Ser Glu Ile
            260                 265                 270

ata cca gca gct att gac gat tac aat gtc caa gca tac att ttc aaa   864
Ile Pro Ala Ala Ile Asp Asp Tyr Asn Val Gln Ala Tyr Ile Phe Lys
        275                 280                 285

gac tat tgg gaa gac att gga aca att aaa tcg ttt tat aat gct agc   912
```

```
Asp Tyr Trp Glu Asp Ile Gly Thr Ile Lys Ser Phe Tyr Asn Ala Ser
    290                 295                 300

ttg gca ctc aca caa gag ttt cca gag ttc caa ttt tac gat cca aaa   960
Leu Ala Leu Thr Gln Glu Phe Pro Glu Phe Gln Phe Tyr Asp Pro Lys
305                 310                 315                 320

aca cct ttt tac aca tct cct agg ttc ctt cca cca acc aag ata gac   1008
Thr Pro Phe Tyr Thr Ser Pro Arg Phe Leu Pro Pro Thr Lys Ile Asp
                325                 330                 335

aat tgc aag att aag gat gcc ata atc tct cat gga tgt ttc ttg cga   1056
Asn Cys Lys Ile Lys Asp Ala Ile Ile Ser His Gly Cys Phe Leu Arg
            340                 345                 350

gat tgt tct gtg gaa cac tcc ata gtg ggt gaa aga tcg cgc tta gat   1104
Asp Cys Ser Val Glu His Ser Ile Val Gly Glu Arg Ser Arg Leu Asp
        355                 360                 365

tgt ggt gtt gaa ctg aag gat act ttc atg atg gga gca gac tac tac   1152
Cys Gly Val Glu Leu Lys Asp Thr Phe Met Met Gly Ala Asp Tyr Tyr
    370                 375                 380

caa aca gaa tct gag att gcc tcc ctg tta gca gag ggg aaa gta ccg   1200
Gln Thr Glu Ser Glu Ile Ala Ser Leu Leu Ala Glu Gly Lys Val Pro
385                 390                 395                 400

att gga att ggg gaa aat aca aaa ata agg aaa tgt atc att gac aag   1248
Ile Gly Ile Gly Glu Asn Thr Lys Ile Arg Lys Cys Ile Ile Asp Lys
                405                 410                 415

aac gca aag ata gga aag aat gtt tca atc ata aat aaa gac ggt gtt   1296
Asn Ala Lys Ile Gly Lys Asn Val Ser Ile Ile Asn Lys Asp Gly Val
            420                 425                 430

caa gag gca gac cga cca gag gaa gga ttc tac ata cga tca ggg ata   1344
Gln Glu Ala Asp Arg Pro Glu Glu Gly Phe Tyr Ile Arg Ser Gly Ile
        435                 440                 445

atc att ata tta gag aaa gcc aca att aga gat gga aca gtc ata tga   1392
Ile Ile Ile Leu Glu Lys Ala Thr Ile Arg Asp Gly Thr Val Ile
    450                 455                 460
```

<210> 12
<211> 463
<212> PRT
<213> Solanum tuberosum

<400> 12
Met Ala Ser Val Ile Thr Thr Glu Asn Asp Thr Gln Thr Val Phe Val
1 5 10 15

Asp Met Pro Arg Leu Glu Arg Arg Arg Ala Asn Pro Lys Asp Val Ala
20 25 30

Ala Val Ile Leu Gly Gly Gly Glu Gly Thr Lys Leu Phe Pro Leu Thr
35 40 45

Ser Arg Thr Ala Thr Leu Ala Val Pro Val Gly Gly Cys Tyr Arg Leu
50 55 60

Ile Asp Ile Leu Met Ser Asn Cys Ile Asn Ser Ala Ile Asn Lys Ile
65 70 75 80

Phe Val Leu Thr Gln Tyr Asn Ser Ala Pro Leu Asn Arg His Ile Ala
85 90 95

Arg Thr Tyr Phe Gly Asn Gly Val Ser Phe Gly Asp Gly Phe Val Glu
100 105 110

Val Leu Ala Ala Thr Gln Thr Pro Gly Glu Ala Gly Lys Lys Trp Phe
115 120 125

Gln Gly Thr Ala Asp Ala Val Arg Lys Phe Ile Trp Val Phe Glu Asp
130 135 140

Ala Lys Asn Lys Asn Ile Glu Asn Ile Val Val Leu Ser Gly Asp His
145 150 155 160

Leu Tyr Arg Met Asp Tyr Met Glu Leu Val Gln Asn His Ile Asp Arg
165 170 175

Asn Ala Asp Ile Thr Leu Ser Cys Ala Pro Ala Glu Asp Ser Arg Ala
180 185 190

Ser Asp Phe Gly Leu Val Lys Ile Asp Ser Arg Gly Arg Val Val Gln
195 200 205

Phe Ala Glu Lys Pro Lys Gly Phe Asp Leu Lys Ala Met Gln Val Asp
210 215 220

Thr Thr Leu Val Gly Leu Ser Pro Gln Asp Ala Lys Lys Ser Pro Tyr
225 230 235 240

Ile Ala Ser Met Gly Val Tyr Val Phe Lys Thr Asp Val Leu Leu Lys

```
                245                 250                 255

Leu Leu Lys Trp Ser Tyr Pro Thr Ser Asn Asp Phe Gly Ser Glu Ile
            260                 265                 270

Ile Pro Ala Ala Ile Asp Asp Tyr Asn Val Gln Ala Tyr Ile Phe Lys
        275                 280                 285

Asp Tyr Trp Glu Asp Ile Gly Thr Ile Lys Ser Phe Tyr Asn Ala Ser
    290                 295                 300

Leu Ala Leu Thr Gln Glu Phe Pro Glu Phe Gln Phe Tyr Asp Pro Lys
305                 310                 315                 320

Thr Pro Phe Tyr Thr Ser Pro Arg Phe Leu Pro Pro Thr Lys Ile Asp
                325                 330                 335

Asn Cys Lys Ile Lys Asp Ala Ile Ile Ser His Gly Cys Phe Leu Arg
            340                 345                 350

Asp Cys Ser Val Glu His Ser Ile Val Gly Glu Arg Ser Arg Leu Asp
        355                 360                 365

Cys Gly Val Glu Leu Lys Asp Thr Phe Met Met Gly Ala Asp Tyr Tyr
    370                 375                 380

Gln Thr Glu Ser Glu.Ile Ala Ser Leu Leu Ala Glu Gly Lys Val Pro
385                 390                 395                 400

Ile Gly Ile Gly Glu Asn Thr Lys Ile Arg Lys Cys Ile Ile Asp Lys
                405                 410                 415

Asn Ala Lys Ile Gly Lys Asn Val Ser Ile Ile Asn Lys Asp Gly Val
            420                 425                 430

Gln Glu Ala Asp Arg Pro Glu Glu Gly Phe Tyr Ile Arg Ser Gly Ile
        435                 440                 445

Ile Ile Ile Leu Glu Lys Ala Thr Ile Arg Asp Gly Thr Val Ile
    450                 455                 460
```

<210> 13
<211> 16
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Upstream primer that is useful for introducing double termination codons in ADPG-PP small subunit cDNA sequence. Includes a KpnI site

<400> 13
gatattggta ccattg 16

<210> 14
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Downstream primer for introducing double termination codons in an ADPG-PP small subunit cDNA sequence. The primer includes a SacI site.

<400> 14
gggggaattc gagctctatc agatgatgat tccac 35

<210> 15
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Upstream primer useful for introducing double termination codons in an ADPG-PP large subunit cDNA sequence. Includes an NheI site.

<400> 15
cccggtaccg ttttataatg ctagcttggc 30

<210> 16
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Downstream primer useful for introducing double termination codons in an ADPG-PP large subunit cDNA sequence.

Includes a SacI site.

<400> 16
ggggggaattc gagctctatc agatgatgat tccac 35

<210> 17
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Nucleotide sequence for an upstream primer useful for restoring the N-terminus of an ADPG-PP small subunit expression plasmid. The primer includes an NcoI site.

<400> 17
gggtcgccca tggctgtttc tgattcg 27

<210> 18
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Downstream primer useful for restoring the N-terminus of an ADPG-PP small subunit expression plasmid. Includes a KpnI site.

<400> 18
ggggcttcaa tggtaccaat atc 23

<210> 19
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Upstream primer useful for deleting DNA sequences that code for 17 amino acids at the N-terminus of an ADPG-PP large subunit. The primer includes an NcoI site.

<400> 19
ggggccatgg cacgtcttga gagacg 26

<210> 20
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Downstream primer useful for deleting DNA sequences that code for 17 amino acids at the N-terminus of an ADPG-PP large subunit. The primer includes an NheI site.

<400> 20
gccaagctag cattataaaa cggtaccggg g 31

<210> 21
<211> 47
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Upstream primer for amplifying the Arabidopsis ribulose bisphosphate small subunit (ats1A) promoter and transit leader coding sequences.

<400> 21
gctcgagtct agaggatccg tggtcgagat tgtgtattat tctttag 47

<210> 22
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Downstream primer for amplifying the Arabidopsis ribulose bisphosphate small subunit (ats1A) promoter and transit leader coding sequences.

<400> 22
cgagctcgcc atggcagtta actcttccgc cgttgcttg 39